CHARACTER AND PERSONALITY

An International Psychological Quarterly

Vol. IV SEPTEMBER, 1935 No. 1

CONTENTS

PAGE

P. E. VERNON (The Maudsley Hospital, London):
"Can the 'Total Personality' Be Studied Objectively?" 1

CHARLES E. SPEARMAN (University of London):
"The Old and the Young Sciences of Character" 11

W. STEPHENSON (University College, London):
"Correlating Persons instead of Tests" 17

CONSTANCE SIMMINS (Institute of Medical Psychology, London):
"Mental Incapacity: The Intelligence of Patients in Mental Hospitals" 25

L. G. STUDMAN (University of London):
"The Factor Theory in the Field of Personality" 34

W. STEPHENSON (University College, London):
"Perseveration and Character" 44

A. A. ROBACK (Cambridge, Massachusetts):
"Race and Mode of Expression: A Preliminary Investigation in Collective Personality" 53

HAROLD D. CARTER (Institute of Child Welfare, University of California):
"Twin-Similarities in Emotional Traits" 61

News and Notes 79
Critical Comment 79
U. S. A. 80
Great Britain 82
Scandinavia 83

Book Reviews 85

Books Received 95

Duke University Publications

The Manic-Depressive Psychosis

By HELGE LUNDHOLM,
Professor of Psychology, Duke University
Paper bound. $1.00

Schizophrenia

By HELGE LUNDHOLM,
Professor of Psychology, Duke University
Paper bound. $1.00

Conation and Our Conscious Life

By HELGE LUNDHOLM,
Professor of Psychology, Duke University
Paper bound. $1.00

Methods in Sociology—A Critical Study

By CHARLES A. ELLWOOD, $1.50
Professor of Sociology, Duke University

Pernicious Anemia

By BEAUMONT S. CORNELL, $4.00
A complete review of the literature, coupled with his own clinical and laboratory studies on anemia.

The American Peace Crusade, 1815-1860

By MERLE EUGENE CURTI, $3.50
A dispassionate account of the successes and failures of America's pioneers in organizations for proscribing war.

City Bosses in the United States

By HAROLD ZINK, $4.00
The interesting details of personal history and political career which characterize this study of twenty city bosses illuminate certain aspects of city government in some of our leading municipalities.

Catalog on Request

DUKE UNIVERSITY PRESS DURHAM, N. C.

CAN THE "TOTAL PERSONALITY" BE STUDIED OBJECTIVELY?*

P. E. VERNON

The Maudsley Hospital, London

In spite of the wide diversity of theories of personality a large number of psychologists might subscribe to certain general principles which will be enunciated, rather dogmatically, as follows: An individual's personality is a totality which is more than the sum of its parts. The mere enumeration of a person's traits and habits does not give us the person himself, since it omits the essential aspect of organized structure. Each single characteristic has to be considered in relation to the whole; for example, the radicalism of an introverted individual is different from an equal degree of radicalism in an extroverted individual, for each trait is dependent on, and modified by, the balance of all the other traits. The structure of personality would seem to be hierarchical; in other words, certain general emotional traits, conative drives and values, and the underlying mechanisms studied by the psychopathologist, are fundamental; and these appear to control and guide the more specific characteristics such as simple habits of behavior.

Such principles are probably implicit in the layman's commonsense conceptions of personality, and they have often been explicitly formulated by various German psychologists. Yet we have to admit that our conviction as to their truth is little more than a conviction. They have not hitherto received much support from experimental scientific investigations of personality. On the contrary, many American investigators, who have devised ingenious objective tests in their attempts to measure personality traits, have found that two or more tests, or sets of ratings, of the same trait seldom intercorrelate at all highly. An individual may be very persistent according to one test, but lacking in persistence according to another.

* This article was written during the author's tenure of the Pinsent-Darwin Studentship in Mental Pathology.

Hence they conclude that personality is in general unorganized and unintegrated, and that there is no such thing as a general and consistent trait of persistence. Instead, their evidence suggests that there are a large number of separate habits of behavior, specific to each situation, and that personality is the aggregate of all these habits.

It is difficult to accept such conclusions as sound, even though they are based on scientific data; hence there must be a flaw somewhere. In the view of the present writer, the flaw arises mainly from the application of unsuitable methods to psychological material, methods which have been taken over directly from the physical sciences. Lengths, weights, and times can be measured separately in terms of absolute scientific units; they are unidimensional variables. The ordinary psychological test attempts to grade traits and aptitudes in the same manner, to turn them into quantitative variables. In the field of cognitive abilities the attempt has been conspicuously successful, but not in the field of conative traits or emotional tendencies. It is true that we often talk of one man as more persistent than another. But their persistences do not differ merely in quantity, they seem to be qualitatively unlike, and they cannot properly be compared (like lengths or weights) because their significance depends so largely on the other outstanding features in each man's personality. To measure a separate trait we have to isolate it from the rest of the personality structure and to assume that two entirely dissimilar individuals who happen to obtain the same score on our tests of persistence really do possess the same trait. We cannot even measure a trait in a single individual at all, because no absolute measuring rods are available for psychological material; we can only compare his behavior with the behavior of a group of other individuals under the same conditions, and we can then say that his reactions to our tests are more persistent than the reactions of 90 per cent or 50 per cent, etc. of people in general. Clearly then our attempts at measurement, whether with tests or with ratings, inevitably disrupt the personality into such separate bits as can be handled by our quantitative techniques, but they cannot put it together again.[1] It is thus only natural that they should lead to the theory that personality consists of such bits.

[1] Professor Spearman has recently commented on a previous presentation of this viewpoint by the writer (*British Journal of Psychology*, 1934, XXIV, 403-407). He holds that correlational statistics, particularly those that deal with multi-dimensional variables, are essentially synthetic, not analytic. Surely, however, statistical

This dilemma is fundamental. We see the individual personality as an organized structure, but have to acknowledge that our notions of it are subjective and unscientific. We try to investigate it scientifically, but its essential nature disappears in the process, and we are left with comparatively meaningless strings of test-scores and standard deviations, and correlations between the scores of large groups of persons. A somewhat similar conflict exists (or existed) in the field of perception. When the contents of consciousness were analyzed down into sensations, images, and feeling elements, the natural percept itself was destroyed. But Gestalt psychologists and others discovered that it was possible to carry out scientific experiments on perceptual wholes, and we ought to be able to apply their outlook and techniques to personality. We must admit also that the whole conception of the total personality and its structure is at present very indefinite. The same was true of perceptual organization and Gestalt; but with the development of more appropriate methods of research, they are becoming progressively clarified. We cannot expect to achieve a systematic view of personality until scientific investigation is able to contribute more to our knowledge of it than it has done so far.

The need for such a development is all the more urgent because many methods of practical psychodiagnosis are not analytic, and no one has yet been able to validate them adequately by means of correlations. The impressions that we derive from observing an individual's features, his voice, gestures, style of writing or drawing, etc., do not ordinarily consist of the individual's standing on isolated traits. We know that the total personality is to some extent manifested in these various modes of expression, and though we must admit that such knowledge is often biased and partial, yet we make considerable use of it in daily life. Again when we deduce a child's temperamental make-up from his manner of performance at Binet or other tests of abilities, or try to sum up his personality on the basis of a free interview, it is a structured and organized conception

---

synthesis is additive, or at least algebraic. Schematically expressed, statistics can examine the relationship between $a \times A + b \times B \ldots$ and $m \times M + n \times N \ldots$ where A, B, M, etc. are psychological variables, a, b, m, etc. are their weights in a group of Subjects. But to examine the relationship of $aAbB \ldots$ with $mMnN \ldots$ in one Subject, and of $a'Ab'B \ldots$ with $m'Mn'N \ldots$ in another Subject, where $aAbB \ldots$ etc. represent organized products, not aggregates, would seem to be outside the present scope of correlations. With the rest of Professor Spearman's article, i.e., with his plea for the combination instead of the opposition of "psychometric" and "intuitionist" approaches, the present writer is in entire agreement.

at which we arrive, which is distorted and destroyed if it is merely recorded in terms of numerical ratings. Graphological diagnoses and the Rorschach inkblot test similarly try to portray personality as an organized whole, so that they cannot be validated (or at least have not yet been validated) by correlations between separate variables. Yet the typical method of the scientific experimenter is, say, to measure the height of the forehead, or the size of specific handwriting details, or to get a number of judges to arrange photographs in order for intelligence. These abstracted measurements are then compared with tests of intelligence or with ratings or tests of other traits, and, as could only be expected, negligibly small correlations are generally found. Hence such experimenters fallaciously conclude that our knowledge of others, based on their expressive movements, or our intuitions of their personalities, are largely illusory. But how can we prove that the study of the total personality, in particular of its general traits and underlying mechanisms rather than of its specific manifestations, is going to yield more meaningful psychological results? One or two interesting, though tentative, experiments along these lines may be briefly described.

A pioneer investigation was carried out by von Bracken at Jena ten years ago.[2] He prepared two descriptions of the personalities of each of 39 children. One description was "schematic," consisting simply of a list of some 26 questions about the personal characteristics of the child with brief answers—generally "Yes" or "No." These descriptions were typical school reports, in which each trait was treated separately. The other descriptions were written without any restrictions, and attempted to portray the most significant characteristics and the structure of the personality. Several adult judges, who did not know the children, were then given one of these descriptions to read, and afterwards had to answer various questions about the child's traits (including traits not actually mentioned in either description), or to deduce which traits were dominant, or to write another sketch of the personality either directly after the reading or quarter of an hour later. A large number of deductions were made, and they were scored by reference to the original descriptions, or their correctness was estimated by the authors of the descriptions who knew the children intimately. In every experiment it was found

[2] H. von Bracken, "Persönlichkeitserfassung auf Grund von Persönlichkeitsbeschreibungen," *Jenaer Beiträge zu Jugend- und Erziehungs-Psychologie*, 1925, I, 1-50.

that the deductions from the free descriptions were more accurate, more comprehensive and more pregnant than the deductions from the schematic descriptions. The judges obtained a much better idea of the personality from the generalized sketch than from the long list of separate characteristics, and according to their introspections they found it easier to comprehend or grasp the personality from the former. It should be noted that the judges always got their impressions secondhand in these experiments. Another shortcoming was that much of the scoring of their deductions was necessarily subjective. Nevertheless, the results are very suggestive of the superiority of "whole" to "part" methods in the study of personality.

Somewhat similar experiments were devised independently by Cantril at Harvard, in 1932.[3] He prepared three descriptions or sketches of the personalities of each of twelve individuals; some of the sketches were couched wholly in general terms, giving the person's most significant traits, while others consisted wholly of objective descriptions of specific items of behavior, manifestations of these traits. On presenting the sketches to twenty judges, it was found that they could usually obtain clearer and more comprehensive impressions of the personalities from the general than from the specific descriptions, since the latter usually failed to cohere together into pictures of "real persons." The judges generally assessed the former more highly than the latter for "trenchancy," i.e., the degree in which "they felt acquainted" with the person described. Moreover, when six of the judges were asked to recall all they could about the personalities three months later, 73 per cent of their recollections were of general impressions (and these were all correct), whereas only 27 per cent of specific items were remembered (some of them incorrectly). As von Bracken has pointed out, memorizability is rather a good criterion of what is essential or significant in a personality. Though this experiment did not compare the structured personality with its separate components, yet it showed that the judges always tried spontaneously to organize their impressions into a structured whole or Gestalt, and that they were able to do so more successfully and more lastingly when they were told the "higher level" or more fundamental traits than when they were told the "lower level" specific actions.

[3] H. Cantril, "General and Specific Attitudes," *Psychological Monographs*, 1932, XLII, No. 5, Chaps. IV and VII.

Neither of these two investigations so far satisfies our search for scientific treatment of the total personality. Descriptions of personalities as wholes could only be validated, or assessed for "trenchancy," by subjective methods; when objectively validated they had to be split up and considered piecemeal. It is doubtful, then, whether any ordinary method of psychological measurement can be applied to the total personality. Yet there is an indirect method which, although it still distorts the personality to some extent, at least recognizes and takes its organized structure into account. This is the matching method *(Zuordnungsmethode)*, which was employed by von Bracken in the course of the research we have described, but whose development has been associated chiefly with the names of Arnheim and Wolff. It was used also by Cantril and Rand in a recent study published in this journal; hence a short summary of it will suffice here.

Suppose we wish to investigate the relation between facial expression and the total personality. Descriptive sketches of the personalities of, say, five subjects are written by close acquaintances; these sketches, together with photographs of five subjects, are then presented to a group of judges. The sketches and photographs are arranged in a different order, and each judge tries to fit one series to the other, to say which sketch matches or goes with each photograph. The number of correct judgments or matchings may be expressed as a contingency coefficient by means of a technique which the present writer has recently worked out. If the validity of the matching, or the relationship between photographs and sketches, is very low, the contingency coefficient is little better than zero; whereas if the validity is nearly perfect the contingency coefficient (like a correlation coefficient) approximates to 1.0. Not only then can the relation between features and personality in this group, or in other small groups of subjects, be formulated quantitatively, but also a separate coefficient may be computed for each individual whose sketch and photograph are compared.

Although a matching experiment conforms as fully as a correlation experiment to the requirements of scientific objectivity, and although quantitative statistical controls are just as applicable, yet the matching method does enable us to handle organized wholes such as the total balance of features, or the pattern of handwriting, or

literary style, etc., and the total personality, without disrupting them into artificial fragments. Much work has still to be done before the method achieves its maximum usefulness, and many cautions have to be observed in applying it (a full account of these points, together with the statistical technique, will be published elsewhere). Nevertheless it appears to provide one solution of the conflict which we have mentioned between psychological theory and objective experiment. We admitted that it is an indirect method, for it still does not compare the photograph with the total personality as such, but only compares the judges' impressions of the one with impressions of the other. This limitation may, however, not be so serious as it appears at first sight. In the view of the present writer, personality has no real existence apart from subjective impressions of external observers; hence it never will be accessible to direct study by objective methods.[4] But this is an epistemological question whose discussion would lead us too far afield; moreover, its outcome does not affect the principles which we put forward at the beginning of this paper. Let us turn then to certain matching experiments which tend to provide objective confirmation of these principles.

I. As part of a wider investigation of personality, 20 to 25 subjects, American college students, were observed by the writer and by two other experimenters while they were engaged on various performance tests.[5] These three experimenters, who either observed separately or in pairs, embodied their impressions of the subjects' personalities in interpretative sketches, and in ratings upon four traits—"intelligence," "quickness," "extroversion," and "emotional stability." When the ratings by the different experimenters were inter-correlated, the average of the twelve coefficients was $+ 0.44 \pm .032$. In addition, the experimenters attempted to identify or match one another's sketches with their own impressions, and were successful in 23.3 per cent of cases. It would have been better if the sketches had been split into smaller groups and matched, say, 5 at a time; 20 or 25 are far too many to deal with conveniently. Thus a greater proportion of correct matchings might have been

[4] See P. E. Vernon, "The Biosocial Nature of the Personality Trait," *Psychological Review*, 1933, XL, 533-548.

[5] The tests are described in an article, "Tests of Temperament and Personality," *British Journal of Psychology*, 1929, XX, 97-117. This article deals with an investigation carried out in England, but the same tests were used in the research at Yale University.

obtained had the technique been more fully worked out at the time of this experiment (1930). Nevertheless the contingency coefficient which corresponds to the 23.3 per cent of successful matchings is $0.72 \pm .055$. This indicates that the experimenters agreed more closely in their general structured impressions of the subjects than in their ratings on separate traits.[6]

II. Two experiments were carried out recently with photographs of 5 of the above subjects whose personalities were extremely distinctive or heterogeneous, and with photographs of 10 more of them whose personalities were much less outstanding and more homogeneous. The 5 or 10 photographs were matched with brief character sketches and were, in addition, ranked on six separate traits—"intelligence, sociability, emotionality, artisticness, leadership and efficiency." A group of 25 judges ranked the 5 and matched the 10 photographs at one experimental session, and matched the 5 and ranked the 10 at another session a week later, or vice versa.[7] The "true" standing of the subjects on these traits had been previously determined as objectively as possible by giving to them a large number of standardized personality and aptitude tests and by obtaining full ratings from various groups of acquaintances. The character sketches were composed on the basis of the subjects' objective test-scores and of the writer's personal acquaintanceship. An attempt was made to limit some of the character sketches solely to the epithets employed in defining the six traits already mentioned, so that a direct comparison might be made between the ranking and matching of photographs against identical criteria of personality. But since these limited sketches gave much the same results as others

[6] Though the writer has found that matching contingency coefficients are usually comparable in magnitude to correlations, and that the probable error of a difference method can legitimately be applied to two contingency coefficients, yet a direct comparison of these rating and matching results is obviously statistically unsound. A possible, though not very satisfactory, method is to express the ratings in terms of a matching contingency coefficient; its value is $0.27 \pm .041$. The difference between this figure and $0.72 \pm .055$ is $6\frac{1}{2}$ times the probable error of the difference, and is therefore significant.

[7] The judges were convalescent patients at a voluntary mental hospital. Care was taken to exclude any patients whose mental disorder prevented coöperation, and there is no reason to regard them (for the purpose of the experiment) as other than an ordinary group of adults. Fifty judges in all took part, 30 male and 20 female, but it was found that the less intelligent were often unable to complete the experiments, and that their judgments were too inaccurate to yield definite results. Hence only the results for the more intelligent half are included here. The writer is grateful to Drs. Slater, Anderson, Lewis, and Maclay, and to the ward nurses for their help in procuring the judges. The same persons undertook the classification of the judges in respect of mental ability.

sketches of each subject written without any restrictions, the results from all the sketches are combined in the following table.

Experiment	*Coefficient and probable error*
Matching 5 photographs with sketches by 25 judges	$C = 0.652 \pm .035$
Ranking 5 photographs on 6 traits (average of 150 coefficients)	$r = +0.244 \pm .024$
Matching 10 photographs with sketches by 25 judges	$C = 0.245 \pm .049$
Ranking 10 photographs on 6 traits (average of 150 coefficients)	$r = +0.044 \pm .018$

It will be seen that the 5 photographs were matched with the sketches much more accurately than they are ranked on the separate traits. In the experiment with 10 photographs the matching coefficient is also superior to the ranking coefficient, though probably the difference is only of moderate statistical significance.[8] But, as already stated, this group of subjects was exceptionally homogeneous, and the judgment of their personalities by *any* method was probably too difficult to yield satisfactory results.

III. In both the above investigations we are hampered by our inability to institute a direct statistical comparison between correlation and contingency coefficients. But one further experiment may be mentioned which avoids this difficulty, since it can be formulated in contingencies throughout. Allport and Cantril studied the relation of the voice both to certain separate components of personality and to personality considered as a whole.[9] They obtained large groups of judges (average number 63) who matched the voices of hidden speakers with objective criteria such as age, height, vocation, and scores on tests for ascendance, extroversion, and values. In 30 of these experiments with separate personality characteristics, 15 different speakers took part, three at a time. On an average 45.5 per cent of the matchings were correct. The present writer has computed the validity of these results in terms of contingencies and finds an average coefficient of $0.25 \pm .009$. Allport and Cantril then combined these same characteristics into summary sketches of the personalities of the speakers, which were matched with the voices by the same judges in 6 experiments. Here 54.5 per cent of the matchings were correct, and the average contingency is distinctly

[8] If the rankings are again expressed as matching, the value of the contingency coefficient is $0.004 \pm .017$. The difference between this figure and $0.245 \pm .049$ is 4.8 times its probable error.

[9] G. W. Allport and H. Cantril, "Judging Personality from the Voice," *Journal of Social Psychology*, 1934, V, 37-55.

higher, $0.41 \pm .020$. The difference between the two results shows a high degree of statistical significance.[10] In other words, the mere fact of grouping the isolated characteristics into a picture of each individual as a whole definitely improved the validity.

These three experiments unanimously show that personality can be more accurately and consistently judged as a structured whole (either from the manner of doing performance tests, or from photographs of the features, or from the voice) than can the separate traits or components into which it may be analyzed. Other experiments along these lines are in progress, for it is obvious that many improvements could be effected in their technique. In conclusion, it is certainly not claimed that the matching method is the only, or even the best, method of studying personality objectively. Nor is it denied that correlational methods might be adapted so as to yield somewhat better results than were given by the separate ratings in the above experiments. But matching does seem to offer a means of approach to a great many problems in the field of personality which are either inaccessible to, or else have hitherto been neglected or misapprehended by, the standard methods of psychological and statistical research.

[10] The difference is 7.4 times its probable error.

THE OLD AND THE YOUNG SCIENCES OF CHARACTER

CHARLES E. SPEARMAN

University of London

Present Block in Psychology

It is generally agreed that nowadays psychology has arrived at a very undesirable degree of disunitedness. Each school, if not each individual, seeks to establish the science independently both of his predecessors and even of his colleagues. The result is that all alike have come into general discredit. Psychology is a byword of reproach among other sciences.

And in no region of psychology has this uncoöperativeness, together with its inevitably concomital weakness, been more flagrantly manifested than in that which deals with individual differences of mental constitution, or, in a word, with personality.

Here the main root of the trouble would appear to lie in the method of investigation. There are two radically different procedures at issue. And each is accused of incompetence by those who employ the other.

In such a painful and sterile situation it becomes eminently desirable that the contending parties should try to get together and submit the trouble to common and mutual helpful consideration. And the present number of this journal would seem to proffer a very favorable opportunity. For it has been fortunate enough to publish the preceding article by Dr. Vernon, which interestingly portrays one of the rival doctrines of method. In contrast, the present writer will endeavor to set forth the case of its opponent.

Old Doctrine of "Faculties" and "Types"

The most ancient and still the most prevalent method of approaching the problems of personality is to take mental constitution as dominated by a small number of "faculties," "powers," or other

traits, including such as constitute "types." In earlier literature the most widely current of these traits were general abilities, such as memory and imagination. Recently a larger rôle has been conceded to general emotional and conative tendencies. Thus Dr. Vernon writes: " . . . certain general emotional traits, conative drives, and values are fundamental . . . and guide the more specific characteristics." As an instance of such guiding traits in highest repute at the present day may be cited that of "introversion."

Protest of Independent Elements

But every now and again—notably by Herbart—such general traits have been sharply challenged. Thus Herbart declares:

> Memory and imagination agreed in that their superior strength is usually limited in every man to particular kinds of objects. . . . He who easily remembers the technical expressions of a science that interests him often has a bad memory for the novelties of town.

Volkmann, following him, declared that every idea has a memory and imagination of its own. In modern times no less emphatic have been the writings of Thorndike:

> The science of education should at once rid itself of its conception of the mind as a sort of machine, different parts of which sense, perceive, discriminate, imagine, remember, conceive, associate, reason about, desire, choose, form habits, attend to. . . . There are only the particular connections between particular mental events and others.

Here, then, the personality is no longer taken to be manifested in a small number of very general traits, but in an extremely large number of very specific ones.

Invention of Correlational Coefficients

As is plain, the fundamental point at issue was as to whether the traits can legitimately be treated as functional unities, or instead are merely classifications. Do, for instance, the different forms or qualities of introversion tend to go together in the same individual? Or, on the contrary, does the presence of one of the forms argue nothing for the presence or absence of the others? The character called into question, then, is that of co-existence (including covariance). Which contending party was right?

To answer at last this old and seemingly insoluble question, Galton came forward towards the end of last century with his invention of a numerical "coefficient," whereby the degree of the disputed

co-existence could be exactly measured. His own particular coefficient was soon replaced by very numerous others; there are now in the field coefficients of "association," those of "colligation," and those of "contingency"; there is the correlation ratio method, the bi-serial method, the tetrachoric method, and many more. They all, however, serve the same general purpose: that of measuring the companionship of different variables (though the range of variation may be nothing more than that between occurrence and non-occurrence).

Now the effect of this new statistical invention upon the ancient doctrine of faculties was disastrous. For the most part, the supposed general traits proved to be illusory. A person might possess one form of a trait in high degree, but another form hardly at all.

Still the protesting theory of independent elements was not satisfied either. Although the different forms were, as just said, not completely linked together, none the less they usually were so in some degree. The new correlational coefficients, though not rising to perfect unity, were nevertheless often much above zero. This situation was still unsatisfactory. Science aims at expressing phenomena in laws that, so far as they go, involve perfect correlation.

Progress to "Factors"

In this grave trouble the idea was suggested that the unities or perfect correlations which could not be found in actual observation might perhaps be inferred hypothetically.

The first application of this idea was momentous. It concerned mental abilities. The coefficients of correlation between these were found to be such that each of them could be expressed as the sum of *two* unities or, as they were called, factors. One of these, *g*, was general. The other, *s*, was specific. About eighteen months after this discovery had been announced, it was verified and utilized by the publication of the famous tests of Binet (in which, however, the term *g* was replaced by "general mental level").

Later on, various further factors were detected, not only in the constitution of ability, but also in that of character. Above all, the latter has been shown to possess a general factor which is designated as *w* and which makes for kindness, unselfishness, conscientiousness, and self-control. This immensely important result, which originally derived from estimates made of character in ordinary life, has now at last been verified by definite observations of conduct under ex-

perimental conditions. Among other factors which have been found to play a very large part in character may be mentioned *p* (a letter chosen in order to indicate its possible kinship to "perseveration").

Under the stimulus of these and many other successes, various further statistical procedures of factorization have been invented; notably, those of Hotelling, Thurstone, Holzinger and Stephenson.

At the present moment—especially in England and America, but increasingly elsewhere also—the problem and procedures of factorization stand in the forefront of psychological research.

Among the outstanding manifestations of this recent activity may be specially mentioned the international committee convoked by Thorndike in order to ascertain whether an individual can be described, with some approach to adequacy and accuracy, by means of measuring a not over-large number of "unity traits," such as the aforesaid factors.

Facts versus Theories

Now for the other side of the question. As said, this new technique—this "psychometry" as it has sometimes been called, especially by its adversaries—amounted to nothing less than a deadly criticism of almost all that had been done by everybody else. Not unnaturally, then, it has met with wide and strong opposition. The resistance to it has taken two ways. The one and less commendable resource has been simply to ignore it in the hope of eventually living it down. The other and more scientific reaction was to accept the challenge, put up a defence, and even launch counter-attacks.

This franker policy to save the old method has found a gallant champion in Dr. Vernon.

To pluck the flower safely out of the nettle danger, he boldly takes up the very case where different tests or ratings of alleged general traits are supposed to have really quite a low inter-correlation. He pleads as follows:

> An individual may be very persistent according to one test, but lacking in persistence according to another. Hence they [the "psychometrists"] conclude that . . . there is no such thing as a general and consistent trait of persistence. . . . It is difficult to accept such conclusions as sound, even though they are based on scientific data; there must be a flaw somewhere.

Surely this comes near the famous saying that if the facts are against the theory, so much the worse for the facts!

Objection to Statistical Technique

However, let us go on to examine the nature of this suggested "flaw." We find that it is partly statistical. The exponents of the new method are charged with endeavoring to measure character in ways which are taken over from the physical sciences. The faults alleged against its variables are those of being "absolute," "unit-dimensional," and merely "additive."

What say the statisticians thus unexpectedly put on their defence? They regard the preceding statistical criticisms as not only unfounded but even a little amateurish. And anyhow they say that —save for one single point—the measurements employed, valid or invalid, are precisely the same with both parties, old and new; both alike employ equalities, inequalities, rankings, marks, estimates, and so forth; both alike may or may not depend upon "intuition" in the sense of general impression. The one and only exceptional point where the two parties do fundamentally diverge is that which we have already seen. It is the case of estimating the degree that any two characters tend to go together. Here, the old party work on general impressions and nothing else. With little if any actual details in their minds, they arrive "intuitively" at the verdict that the correlation between the two characters is "good," "very good," "bad," or "very bad." And it is just intuition of this kind, says the new party, that has shown itself to be grossly misleading; so much so, that the invention of coefficients came as a scientific god-send.

Objection on Psychological Basis

More favorable becomes the position of the old facultist party when it desists from trying to overcome the statisticians on their own ground, and turns instead to that of psychology, where it meets its opponents on more level terms. Or rather, it has a great advantage—that of finding a potent ally. The old theory of faculties is now rejuvenated by fusion with the recent theory of Gestalt. Each single character, Dr. Vernon urges, has to be considered in relation to the *whole*. Whereas the correlational methods, he urges, "inevitably disrupt the personality into such separate bits as can be handled by the quantitative techniques, but they cannot put it together again."

Now, by themselves these words suffer from the vagueness which has been said to permeate the theory of Gestalt. However, our

present author carefully guards against danger of misinterpretation by giving us many examples of what he is approving or condemning, as the case may be. Among the approved traits, we have already mentioned his "general emotional traits, conative drives, and values." To these he adds "the impressions we derive from observing an individual's features, his voice, gesture, style of writing or drawing," "deductions on the basis of a free interview," "graphological diagnoses and the Rorschach inkblot test," "deductions of dominant traits," "sketches of personality." The traits disapproved include: "simple habits of behavior," "sensations, images, and feeling," "isolated traits," "numerical ratings," "height of forehead or the size of specific handwriting details," "arrangement of photographs in order of intelligence," "schematic description," "objective description," "lower level specific actions." But in reply to this separation of traits into sheep and goats, the young statistical party is able to say that its method includes in its scope both the one and the other. Any character whatever, general or specific, complex or simple, structural or unstructural, high level or low level, additive or otherwise—so long as it can occur in varying degree, or in varying quality, or sometimes not at all—still furnishes food for the method of correlations.

And at this point the young party can bring in what is perhaps their strongest point of all. The use of their method does not preclude the employment of any others. This method determines how far different variables tend to go together. Psychological investigation cannot indeed override these results—as the facultists may be accused of attempting to do—but it certainly can and must vastly transcend them. For example, even factorization, as the present writer believes, is of little or no service by itself. Its proper use is as a means of verification. Its results should be such as to suggest hypotheses or to decide between those suggested already.

The net result of all these comparisons between the old method and the young would appear to be comfortable. The fundamental difference between the two has shrunk to one single point, the scientific way of estimating the degree in which two characters or other items tend to go together. Here the old method uses general impressions; the new method, statistics. But on the sole basis of this one fundamental divergence, there do indeed rise up others of amazing magnitude—enough perhaps to transform the whole of psychology.

CORRELATING PERSONS INSTEAD OF TESTS

W. STEPHENSON
Psychological Laboratory, University College, London

I

This is not going to be a statistical article. Instead, I am to introduce a new general technique to readers of this journal, one which should be of greatest interest to them, because its widest use will be in the study of human personality. The technique is a complete inversion of all previous factor techniques.[1] It is a very simple notion indeed, but it brings a new tool to the aid of psychology, general, social, or individual.

It is necessary, first, to give a brief account of the present-day factor technique.

The Present-Day Factor Technique

Today we correlate tests or estimates. The work of Professor Spearman and those he has so greatly influenced in England and America, whether along statistical or experimental lines, has been concerned with group tests and estimates, such as tests of intelligence and estimates of character qualities. These are used with large populations of persons, and correlated.

The correlations for a number of tests, as is well known, are next analyzed in terms of factor theorems, like the Two Factor Theorem of Professor Spearman.

All this serves two ends. The analysis can only be, or should only be, a scientific check or test of psychological hypotheses. All science proceeds along this general line, and psychology cannot be an exception. Next, science is concerned not with unique events (an axiom that some workers in psychology overlook), but with repeatable processes, and correlations are above all things indices of this repeatability.

[1] It was described as such in a letter to *Nature*, June 30, 1935.

Present-day factor work, then, can be truly scientific. Its result is to tell us first how people are universally *alike*, and thereupon to allow us to measure them for the processes in which they are alike, because only with such a warranty could we compare them. Thus we can measure *all* individuals, normal or otherwise, babies or adults, for a few universal factors such as Spearman's *g*, *c*, and *w*, which represent, it so happens, the broadest features respectively of cognitive, affective, and volitional behavior.

Large populations of persons, however, are required in experiments that are made on these factors, and as a consequence the quality of the testing suffers. The size of the population sets a limit on the tests that can be applied, and on the apparatus and controls that can be used. All the refinement and delicacy, the intricacy and subtlety of the laboratory, are lost just when it is most needed in the study of the most complex of all activity, human behavior. Experimental work is slowed down, almost to the point of being burdensome. A research student may spend two years isolating a single factor. All one's controls and hypotheses have to be sandwiched into the one experiment, to be performed on one and the same group of persons. One cannot perform an experiment today and use its results for another tomorrow. Tests cannot be changed in the way that a chemist changes the mixtures in his test-tube to try out his hypotheses. In short, the present-day technique lacks the pliability that the energetic experimentalist wants at his command. It is a device for massive field work, and not for the clinic, the laboratory, or for rapid and subtle experimentation. On the other hand, this massive technique has features that no other can replace.

But the problem arises, and it is one that has bothered me for some time, can the above disabilities be removed? Can we make factor studies on a *few* individuals, and bring the methods of correlations and factor analysis into the laboratory?

The New Technique

Fortunately, for a large part of psychology, these questions can be answered affirmatively. The solution is to correlate *persons* instead of *tests*. In my letter to *Nature* I indicated that all present-day factor theroms could be so inverted.[2]

Whereas previously a large number of people were given a small number of tests, now we give a small number of people a large num-

[2] *Ibid.*

ber of tests or test-items, or require a large number of responses from them. Previously individuals obtained scores; now the tests get them instead, due to the operation of the individuals upon them. By the present-day technique we obtain the factor saturations or loadings of tests, but by the new one we can obtain saturations for individuals. By the older technique we could estimate a person's *g*, or *c*, or *w*, or other factor; now we can obtain factor estimates for test-items.

But a concrete example of the new technique in use will do more than any further description to illustrate some of the above points, and in the following section of this article I describe one such study, made in a few hours, concerned with those highly personal things, our predilections and tastes.

The whole domain of psychology, however, is open for the technique to explore: introspective psychology in general, our preferences, sentiments, motives, and ideals; studies in aesthetics, such as investigations into the nature of beauty; studies in industrial psychology, on fatigue, sensory and motor acuities, and the like; studies in educational psychology, such as examinations of the abilities of teachers to mark essays or drawings. The technique reaches into all these, no less than into those somewhat rarified regions of sensory and psycho-physical psychology.

II

Experiments on Taste or Predilection for Color

We are not, of course, here concerned with those tastes which require the tongue for their perception, but with the predilections that we have for some things rather than for others. The new factor technique described in the preceding short note could be used to conduct experiments on both kinds of tastes, but here I am to be concerned with the latter only.

Tastes and matters of tastefulness make up a good part of the stock-in-trade of the educated person. It is not that the person *likes* one thing more than another, or that it gives him greater pleasure only. I may like old pipes, even to the extent of being sentimental about them, but my taste may be for choice cigars. To have taste is to be discerning, to relish things intellectually no less than affectively or by way of sentiment.

We have to distinguish, too, between tastes and active participation in them. I may have a taste for etchings or for Giotto-like

paintings, for old books or for fashionable suits, and yet neither possess any of these nor take any active interest in them. If I had the opportunity, then no doubt I might possess these things for which I have a taste. Tastes are merely tools to be used when an appropriate occasion arises.

Taste is all a matter of opinion, so popular opinion avers, and one man's taste might be the abomination of another. Tastes are subject to fashion, and are highly partial too. We have tastes for clothes and luxuries generally, but I have yet to hear of the engineer with a taste for A.C. generating motors.

Finally, no one would deny, I hope, that tastes are part of one's personality. Obviously their importance in a thousand and one walks of life, for success in social and commercial activity, is of no little moment, and one feels that psychologists should give them far more of their attention than they have received hitherto. For indeed very little has been done on the scientific study of tastes. All that I have written above may be true of them, but woeful gaps are left to be filled in by experiment. What kinds of tastes are there? How specific in fact are they? Are they connected one with another? Have tastes a pleasure content, or are they intellectualized only? Are the things we like the same as those we have a taste for?

The new factor technique can help us to develop and solve these questions. It is outside my present purpose, of course, to answer them all. I am merely to point to the problems and to illustrate the efficacy of the new technique to solve them. To this end, then, I describe the following short experiment on taste for color.

The Experiment

I took from an artist's colorman a sample of every homogeneous colored paper that he stocked, sixty different colors in all. The sample contained some brilliant poster colors, vivid reds, greens, blues, and oranges; there were pure and delicate blue-green hues and salmon tints, a range of tinted browns, greens, and greys. No blacks or pure greys were included, and no very light creams or whites. All the papers had the same "matt" surface.

As subjects I took 12 young women attending lectures, and 8 men of the same range of age and social status.

Each subject graded the colors for me on a scale ranging from 10 for the color he liked best, to 0 for that he liked least. No reference had to be made to the use to which any color might be

put; judgments had to be taken quickly and as spontaneously as possible. To ease my later calculations, the subjects graded the colors in such a way as to supply a pre-arranged frequency distribution as follows:

Score (x):	0	1	2	3	4	5	6	7	8	9	10
Frequency:	1	2	4	7	10	12	10	7	4	2	1

Thus *one* least pleasing color was given mark 0, the *two* next least got mark 1 each, and so on up the scale.

Each of the sixty colors, then, received a score (x) from each of the twenty subjects. To obtain reliability coefficients the whole grading was repeated.

Now just as would happen if we had 60 people scored on 20 tests, correlations can be calculated for the 60 colors scored by 20 people. In the former case r_{12} would be the correlation between tests No. 1 and No. 2; in the latter case r_{AB} would be the correlation between the scores given to the colors by subjects A and B.

The experiment should help us to answer several questions. If each individual has his or her own specific likes and dislikes, then the individuals will not correlate except by chance alone. But if significant correlations appear, they could be factorized. It might then be possible to describe as separate tastes any unitary factors that emerge, to determine their range, and to measure individuals for them.

The Results

The facts in a summary form are as follows. They, and others like them, are to be published in detail elsewhere.

Twelve individuals, 10 women and 2 men, correlated positively among themselves a mean amount 0.36 ± 0.065, 0.66 being the highest correlation. The rest correlated a mean amount —0.15 with these twelve, and + 0.20 on the average among themselves.

That is, to a first approximation two distinct factors have emerged, one among the twelve individuals, and the other among the remaining eight, the one tending to be the obverse of the other. Incidentally we can see from this example how useful the technique will be to put "type" psychology on a scientific footing.

Next, the factor saturations or loadings of these individuals can be determined, using for convenience the Spearman theorems. The factor among the 12 persons we shall name factor I, and the other factor II. The saturations of the individuals with these factors,

that is, an index of how well they serve to grade the colors for the factor concerned, are as follows:

Subject No.	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Saturation with factor I:	.70	.65	.44	.63	.61	.63	.60	.57	.55	.63
Saturation with factor II:	—	—	—	—	—	—	—	—	—	—
Subject No.	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*
Saturation with factor I:	.68	.72	—	—	—	—	—	—	—	—
Saturation with factor II:	—	—	.50	.44	.32	.53	.44	.34	.43	.55

That is, subject No. 1 had a factor saturation 0.70 for factor I, but none (except obversely) for factor II; and so on along the table.

Finally the factor content of each color can be calculated, just as we can estimate the *g* of an individual by the old technique, now for *either* factor I or for factor II, in terms of the two groups of persons respectively.

Explanation of the Factors

What, then, can be said about these two factors? Factor II is for those persons who rated the pure bright colors as most pleasurable. Factor I, on the other hand, is for those who chose the subtle tones and hues, the subdued colors, as most pleasing. The former colors would be considered to be vivid but crude by a person of taste, while the latter are nowadays regarded as the most tasteful.

Now an advantage of this technique is that there are only a few persons concerned, and I know them all extremely well. Those with factor I are certainly the more artistic and tasteful in general; two are art students, and all dress with taste, although not necessarily "smartly." (This raises the problem, indeed, whether those who dress in ultra-smart fashions are lacking in taste.) Those with factor II are frankly inartistic; some dress crudely; one horrified me with his taste, or lack of it, for furniture; and in all of them the bright colors have appealed to them as they do to children. The exquisite blue of a sugar bag is merely a crude color to these people, while the brilliant vermillion, that shocks the more tasteful person by its very intensity, is highly esteemed. It may well be, too, that women are usually more tasteful for color than men—a matter for no surprise to judge by the clothes that men wear, and the fashions that women have to attend to all their lives.

Thus, not only do the subjects with factor I apparently obtain more pleasure from the colors they all tend to rate alike, but these colors are by modern standards the most tasteful. Pleasure and taste agree in this instance, although, to be sure, experiments can be performed to check up thoroughly the point.

Extension of the Work

Now all this is just a beginning. To extend the work, we could use the same series of colors to compare individuals from different social groups, from different professions and occupations. Children can be compared with adults, and children of one age with those of another. From the factor estimates obtained for the colors new tests can be devised, with more uniform colors for the factors concerned. This, indeed, is one of the great services performed by the new technique of correlating persons, for it allows us to construct tests with *known* factor contents, and these, again, can be applied in due course by the older technique of correlating tests.

Or the range of colors could be extended. One could even devise interesting tests of color-blindness involving the new technique.

Individuals could be asked to select those that they believe to be the most tasteful, instead of merely the most pleasing, to see whether knowledge, pleasure, and taste are associated.

Again, we could correlate the same individuals, or a specially selected group or groups, for other series of tests, designed to investigate other tastes. Tastes for vases, cloth materials, furniture, perfumes, pictures, literature, and so forth can be factorized. It could be determined whether these various taste factors are themselves interconnected. So with intimate control of the experiments a science of tastes in general can be built up.

Nor would this work cease with the discovery of factors and the investigation of their range. Since only a few individuals at a time need be examined, it should be possible to enter into great details, to find out still more about the nature of the various factors. Are there, for instance, deeper unconscious processes underlying our preferences and predilections? Is it intelligence, opportunity, heredity, or environment that determines one's tastes?

Conclusion

Sufficient has been said, I hope, to give an introduction to the scope of the new factor technique. Its use on small numbers of

persons and its amenability to rapid experimentation are noteworthy. I performed the above experiment within an hour and worked out the results in two more. Its use as an essential tool to check the theories of "type" psychologists, and as an aid in the construction of homogeneous tests for further application by the new or the older techniques, as well as its wide range of practical usage, is illustrated in the above experiment and in the extended work that it merits.

Finally, one can factorize the correlations as one finds it convenient or appropriate to do, using Spearman's or anyone else's factor theorems.

MENTAL INCAPACITY:
THE INTELLIGENCE OF PATIENTS IN MENTAL HOSPITALS

CONSTANCE SIMMINS
Institute of Medical Psychology, London

The mental hospital patient differs from the rest of the community in such a marked degree that confinement and close supervision are considered a necessary safeguard. It is a matter of practical importance and scientific interest to discover whether this abnormality of thought and behavior results from a mental equipment differing from that of normal persons, or whether a normal mental equipment is directed to abnormal uses. Are the underlying cognitive and temperamental factors, which determine thought and behavior, blended within the same limits of variation in both the sane and the insane, or are any of these factors present in the insane in an exceptionally high or low degree, and, if so, is this abnormality of mental make-up a condition of the unbalanced mind? Does the insane person fail to make effective, objective use of his abilities, or are the abilities themselves abnormal?

In another article of this journal Studman shows that some types of mental illness appear to be characterized by the presence, in an unusually high or low degree, of certain temperamental traits. It is the aim of the present writer to discuss the cognitive abilities of mental hospital patients, in as far as these can be ascertained by tests of intelligence and of memory.

Not all types of intelligence tests are suitable for use with the insane. Many of the tests that have been used with normal persons consist of words and pictures. If tests of this kind are applied in mental hospitals, it is found that many of the patients respond with a flow of personal experiences, associated with the word or picture presented to them. Their attention is distracted from the task in

hand, and in this way they are hindered from doing satisfactorily tests that are actually within their capacity. The material of the tests needs to be as nearly meaningless as possible, to avoid the disturbing influences of past experience. For the present investigation,[1] non-verbal, visual-perceptual tests, consisting of simple diagrams, were selected as most nearly meeting the special requirements.

These tests, which were scored on a 0 to 20 scale, were given in the first instance to a "control" group of non-psychotic women, selected to represent a very wide range of ability, from university women holding a high honors degree to persons of extremely low intelligence attending an occupation center for the feeble-minded. It was found that no mentally defective woman made a score of more than 7, and no professional worker a score of less than 17, while the scores of the domestic workers ranged from 9 to 16.

The tests were then applied to one hundred female patients at Horton Mental Hospital. Of these, only one made a score as high as 17, although not less than twenty of the patients had at one time been in careers open only to women of superior intelligence. As many as 40 made scores as low as those of the mentally defective non-psychotics, although not more than four were cases of congenital mental defect. In general, then, the test results of these patients seem to suggest impairment or deterioration of intelligence.

An attempt was now made to assess the extent to which deterioration had occurred in each patient tested. To do this it was necessary to obtain information about the patient's intelligence before mental illness set in. Records of school progress and of early employment may provide indications of the patient's earlier capacity, but reliable information of this kind is often lacking. Even when full and accurate records are available, they may be misleading, for types of education and employment are often conditioned by the economic situation and by family tradition, rather than by the capacity of the individual. It was clear that estimates of earlier intelligence based upon case-histories would often be wide of the mark. A more reliable indication of the earlier, undeteriorated intelligence of each patient was obtained by means of the Terman Vocabulary Test. This method had already been used successfully by Babcock, who investigated mental deterioration in patients at the Manhattan

[1] C. Simmins, "Deterioration of 'G' in Psychotic Patients," *Journal of Mental Science*, 1933, LXXIX, 704-734; "The Measurement of Mental Deterioration," *British Journal of Medical Psychology*, 1934, XIV, 113-121.

State Hospital, U. S. A. She found that the vocabulary test gave a useful measure of the amount of intelligence that had been present before impairment began. "Words when once learned are not quickly forgotten, and remain as indications of the ability a person once had."[2] The disparity between the score made by a subject in the vocabulary test and the score made by him in an intelligence test was used by Babcock as a measure of the extent of deterioration associated with the pathological condition. The patients, who suffered from general paralysis of the insane, were tested immediately before "malaria" treatment, and again three months later. Those patients who, according to the medical view, were improving showed at the second testing marked gains in their intelligence test scores, and consequently a smaller degree of impairment as measured by the difference between these scores and the scores for vocabulary. Patients who were not improving showed either no change, or an increase in the amount of impairment measured by the tests.

The picture presented by the vocabulary scores of the Horton patients contrasts strikingly with that of the intelligence test scores. The vocabulary results might be those of a hundred normal persons taken at random from the whole population. The average score is the same as that given by Terman for the normal adult, and the individual variations are not greater than those found among normal persons. It seems justifiable to conclude that, before deterioration set up, the psychotic women did not differ from the rest of the population in respect to intelligence.

The vocabulary test score minus the intelligence test score was taken as indicating the degree of deterioriation of intelligence. The patients were allotted deterioration scores on a 0 to 7 scale, 0 indicating no deterioration, 1 slight deterioration, 2 rather more deterioration measured. There were, however, certain cases in which the vocabulary score could not be accepted as a reliable indication of the amount of earlier intelligence. These were cases of wide-spread mental dilapidation, such as is found in terminal dementia and in some senile patients, where scores were very low in all tests of cognitive abilities, including the vocabulary test. It is probable that the vocabulary scores of patients suffering from general paralysis of the insane cannot be accepted at their face value as measures of earlier intelligence. Both in the vocabulary test and in the test of memory

[2] Babcock, "An Experiment in the Measurement of Mental Deterioration," *Archives of Psychology*, 1930, XVIII, No. 117.

for words, numbers, and pictures, their achievement is lower than that of any other group of patients (with the exception of the epileptics). There seems, therefore, to be some loss of capacity for reproducing past impressions, whether recent or remote. If this is so, the vocabulary score of these patients will under-estimate their earlier capacity, and, consequently, their deterioration scores will under-estimate the amount of impairment that has actually occurred. Neither can the earlier intelligence of epileptics be measured by a vocabulary score, for epileptics appear to suffer from some form of specific verbal disability. The present investigation shows them to be handicapped both in the vocabulary test and also in memory for simple words. There is no evidence yet to show whether this verbal weakness is directly associated with epilepsy, or whether it is a result of the limits set by their illness to attendance at school and to social experience generally.

There remained 70 cases in which it seemed permissible to accept our method of measuring the amount of intelligence present before the onset of illness. Of these, 50 per cent showed deterioration of intelligence. If we add to these the patients who showed not only impaired intelligence but all-round dilapidation, we reach a total of 70 per cent. Impairment was present in the majority of cases in each category tested (mania, melancholia, manic-depression, delusions, dementia-praecox, and general paralysis of the insane).

It is of interest to consider in more detail the 30 per cent whose test scores gave *no* indication of deteriorated intelligence. An analysis shows that this group comprises:

1. The only cases tested who proved eventually to be non-psychotic and who were therefore discharged or transferred.
2. All cases discharged as "recovered" shortly after the date of testing.
3. The only puerperal cases tested.
4. A certain number of the dementia-praecox patients.
5. Six cases rated by the medical officers as being "fairly good" in respect to degree of illness (one paraphrenic, one melancholic, two manic-depressives, and two manics).
6. Two melancholics rated as "medium," one of them "improving," one "fairly bad" case of dementia-paranoides, and one "really bad" case of paranoia.

Subsequently the tests were given to eighty-six more cases. Twenty-six of these were in the Bethlem Royal Hospital, which

admits only such cases as are expected to recover, and the remaining sixty cases were in the Maudsley Hospital, where only voluntary cases are treated. As was to be expected, if our method of measuring deterioration of intelligence was a sound one, both the frequency and the amount of deterioration found were very much less than at Horton, where all the patients were certified as insane. The following table shows the relative positions of the three hospitals in respect to deteriorated intelligence.

Hospital	*No. tested*	*Percentage showing deterioration*	*Average amount of measurable deterioration*
Horton	85	70%	2.44
Bethlem	26	38%	1.85
Maudsley	60	15%	1.29

In addition to the visual-perceptual tests of general ability or intelligence and the Terman Vocabulary Test, there were included tests of memory for recent impressions of words, numbers, and pictures. The tests were scored on a 0 to 20 scale. In the successful performance of these tests two main processes are clearly involved: (1) perception, i.e., the initial awareness of the words, numbers, and pictures; (2) the reproduction of the items thus experienced. It is known that, given adequacy of stimulus and soundness of the sense organs involved, clearness of perception is largely a function of general intelligence,[3] whereas reproduction is an independent process.[4] The act of reproducing past experience under test conditions is conative. Its occurrence is, of course, proof that the impressions gained in the initial experiences have been retained. It was possible by a mathematical procedure to discount the part played by intelligence in the memory tests by deducting from each person's score the part contributed by intelligence. After the memory test scores had been treated in this way, it was found that the scores of the non-psychotic control group ranged from 5 to 13, and that 88 per cent of the patients tested at Horton and 95 per cent of those tested at Bethlem and Maudsley made scores that fell within this range. Two delusional patients and two manics made scores which slightly exceed the highest scores of the non-psychotic group. Of the thirteen patients whose memory scores may be regarded as abnormally low, seven were senile, and two were cases of congenital mental deficiency

[3] Wm. Line, "The Growth of Visual Perception in Young Children," *British Journal of Psychology, Monograph Supplements,* 1931, V, No. 15.

[4] C. Spearman, *The Abilities of Men* (New York and London, Macmillan Co., 1927), Ch. XVI.

with epilepsy. With the remaining four (three delusionals and one dementia-praecox) one is tempted to suspect interference by inhibition or distraction with the conative element in recall. The tests do not, of course, discriminate between ability to *retain* impressions and ability to *recall* what has been retained.

Although only thirteen of the patients tested made memory scores lower than any score of the non-psychotic group, the question arises whether there may not be among the remainder evidence of some slight degree of lessened ability to recall recent impressions, though not sufficiently serious to place the patient outside the normal range. The average score of the non-psychotic group is 9. All categories of patients tested at Maudsley and Bethlem, and three of the categories tested at Horton (delusionals, manic-depressives, and manics) make average scores of approximately 9. The remaining four categories of Horton patients make lower averages:

Dementia-praecox	7.58
Melancholia	7.33
G. P. I.	7.3
Epilepsy	6.5

The low scores of senile patients bring down the average in the dementia praecox and melancholia groups. There is no evidence that the performance of non-senile patients in these groups is poorer on the average than that of the non-psychotics. In the case of patients suffering from G. P. I. (general paralysis) and epilepsy, the lower averages appear to indicate some impairment or disablement (other than deteriorated intelligence), affecting the ability to recall recent impressions as measured by these tests. The relatively low scores of the epileptics are to be accounted for in part by the verbal disability associated with epilepsy, to which we have already made reference. It should be borne in mind that we are here considering *average* results only. There may, in any group, be individual patients whose capacities (apart from intelligence) for recalling recent impressions are somewhat lower than before the onset of their illness. No technique, such as is now being developed, for measuring deterioration of intelligence, has been devised for locating and measuring impairment of other abilities.

The experiment with the memory tests shows that apparent weakening of memory for recent impressions is mainly due to weakening of intelligence, and that, if the intelligence factor is

discounted, the great majority of the patients tested remain, in respect to memory, within the normal range. Similarly, there is every reason to conclude that impaired intelligence affects the general activities of the patients, not only in psychological experiments of every kind, but also in most, if not all, of the ordinary situations of life. It seems reasonable to suppose, too, that this impairment handicaps him in dealing with his own stock of ideas gathered from the whole of his past experience, and is, therefore, one factor accounting for the confused and bizarre notions of many psychotics.

The results of this inquiry into the intelligence of the insane seem on the surface to be fairly definite, and one might be tempted to accept them as a final answer to the question raised in the opening paragraph of this paper. Taking the results at their face value, it appears that in the majority of psychotic patients the mind is a damaged instrument. Intelligence, which is one of the important factors in successful adjustment to life, is, it seems, definitely impaired in the great majority of cases. But an acceptance of this answer to our question depends upon our acceptance of the intelligence test scores as *reliable* measures of the patient's present ability. It may be argued that some part of the patient's ability may be not absent but, as it were, locked up and shut off, completely absorbed in subjective processes, both conscious and subconscious, and that it is, therefore, impossible for him to apply more than a part, perhaps a very small part, of his total capacity to the solving of an intelligence test problem. Our tests may measure only that fraction of his total ability that he is able to direct to effective, objective use.[5] It is true that there is some circumstantial evidence in favor of generally accepting intelligence test scores of abnormal persons as reliable indications of their intelligence. A person suffering from a severe neurosis and entirely failing to make an intelligent adjustment to life may prove on testing to be of superior intelligence. The neurosis, which prevents him from making effective use of his ability in the ordinary affairs of life, does not handicap him in the performance of intelligence tests. Again, a very difficult, maladjusted child with an intelligence quotient of 75 was placed in a new and very favorable environment where the behavior difficulties entirely disappeared. When she was tested again after an interval of two years, the intel-

[5] See Line, Griffin, and Anderson, "The Objective Measurement of Mental Stability," *Journal of Mental Science,* 1935.

ligence quotient was found to be precisely the same as on the previous occasion. The obvious inference is that her intelligence was well and truly measured at the first testing, although at that time emotional disturbances prevented her from using her intelligence effectively in everyday life.

At Horton the intelligence test scores of five patients indicated greater latent intelligence than observation of these patients had led the medical officers to expect. In the remaining cases the test results agreed fairly well with the doctors' estimates. Now this agreement might be regarded, at first sight, as evidence of the reliability of the intelligence tests. But one is forced to admit that the facts bear another interpretation. It may be argued that the tests fail to tap fully the capacity of the patients, since they may be as little able to direct their full capacity to objective uses in the special conditions of the test situation as they are in the ordinary situations of hospital life. The problem remains unsolved. We do not know whether the apparent deterioration of intelligence in a psychotic patient is merely a failure to direct intelligence effectively to objective ends or whether it is in fact diminished intelligence.

The present writer is convinced that the next step towards the solution of the problem depends upon coöperation between psychologists and physiologists. An actual change in the amount of a person's intelligence would presumably be consequent upon a change in the hitherto unknown physiological basis of intelligence. A number of mental hospitals are conducting physiological experiments which may be relevant to our problem. To give one instance, in certain mental hospitals in the United States[6] and in Germany investigations are being made into the passage of substances from the blood into the cerebro-spinal fluid and into the ratio of distribution of these substances in the two fluids. Measurable divergences from the normal ratio were found in 60 per cent of 210 cases of schizophrenia. Nearly all (25 out of 27) of those who recovered or showed definite improvement were found to have distribution ratios within the normal range. May there not be some significance in the fact that approximately the same percentage of schizophrenic patients tested at Horton showed deteriorated intelligence, those shortly afterwards discharged as "recovered" showing no such deterioration? The ground is prepared for a further experiment conducted

[6] Malamud and Rothschild, "The Barrier between the Blood and the Cerebro-Spinal Fluid," *Association for Research in Nervous and Mental Disease*, 1931, X.

jointly by physiologists and psychologists to inquire whether there is any correlation between the passage of substances from the blood into the cerebro-spinal fluid and degree of intelligence. Positive results from such joint experiments as this would throw light upon the basis of intelligence, while negative results would at least clear the ground for other researches.

THE FACTOR THEORY IN THE FIELD OF PERSONALITY

L. G. STUDMAN
University of London

In no field of psychology is it more desirable to come to some common agreement as to the factors involved than in the field of personality, and it seems probable that the factor theory which has proved so successful in unifying and elucidating diverse and often conflicting concepts of cognitive function may be of service here also.

The objection has sometimes been raised that factors are an abstraction and have no place in a dynamic study of personality because any factor studied in isolation is distorted and unreal, in the same way that a portion of a living organism is altered in its very nature if it is dissected. This objection ignores the meaning of the word "factor" which implies a relationship to the whole and to the other factors. It ignores, too, the value of a knowledge of structure in studying function. Thus knowledge of the circulatory system, the central nervous system, and the autonomic nervous systems, would hardly be considered irrelevant in a study of the functioning of the human body, even though these systems cannot be extracted from concomitant matter without loss of function.

Another common objection is that a detailed analysis of factors is valueless in studying personality because each individual is unique; that individual experiences determine reactions to any situation so that every environment becomes a different environment for each person.

Yet these critics usually assume that behavior is to some extent predictable, and, if it is predictable, there must be some element which remains constant for the individual, even though that constant element may be, paradoxically, a tendency to be variable in his behavior. They assume, too, that, one individual may be compared with another, thereby implying some common feature in respect to which

they may be compared, and it is immaterial whether this feature be termed trait, disposition,[1] quality, or factor.[2]

The factor theory does actually provide a basis for discovering what are the really significant units to be observed in the personal make-up, and though our knowledge of these factors may not at present be complete, it seems essential to attempt to establish what are the broad functional unities which may help us to understand the complex interplay of desires, motives, emotional factors, and cognitive ability which constitute personality.

As factor studies are based upon observed facts, measurements of some kind either in the form of records, tests, or estimates are required, and here more objections are raised. Many people protest that such factors are indefinable and too subtle to be observed under experimental conditions; but, while one admits this difficulty,[3] subjective judgments which cannot be verified may be very misleading. For instance, the present writer observed what appeared to be significant differences of behavior in response to the Terman Vocabulary Test. Certain subjects would attempt to define words far beyond their knowledge, while others required constant encouragement to persuade them to attempt words whose meaning they understood perfectly. When, however, this response was observed systematically for all subjects, it was found that such apparently obvious differences were unrelated to general behavior. Many of our judgments on personality rest upon interpretation and evaluation of behavior without any supporting evidence that our interpretation is the correct one, or that the behavior in question has any general significance. Nevertheless, it is of course true that personal observations are extremely valuable and are usually the basis of testing techniques; it is only suggested that in spite of the difficulty of attempting to standardize the conditions of observation, when this can be done the value of such observations is enormously increased.

When the significant factors involved in personality have been discovered and even measured, it remains true that they are of little significance when studied in isolation—it is their relative weight in

[1] See W. Stern, "On the Nature and Structure of Character," *Character and Personality,* 1935, III, 271.

[2] For a fuller discussion of such objections, see C. Spearman, "The Battle between Intuitionists and Psychometrists," *British Journal of Psychology,* 1934, XXIV, 403.

[3] For a summary of the various attempts that have been made to overcome such difficulties, see P. E. Vernon, "The Testing of Intelligence," ed. Professor H. R. Hamley.

the make-up of the individual which is of importance. This balancing or weighing-up of all the factors is more easily achieved at present subjectively than objectively. A trained observer with experience such as the psychiatrist possesses is able to assess many factors almost simultaneously, and his judgment on any one of them is colored by his knowledge of the other factors. He "makes allowances" for differences of age, physique, previous training, intelligence, and so on, and his judgment on any one point is relative to these other factors.

Testing techniques are obviously more cumbersome and such comparative data harder to calculate; but some such relative weighing up is necessary for correct interpretation of results. For example, Dr. May Smith uses a dotting test in which scores seem to depend upon intelligence *and* upon sheer determination, such as is exhibited by obsessional patients. In this case it is clear that the meaning of any particular score can only be discovered if one of these two influences can be measured independently. If the intelligence is already known, the results of the dotting test can be compared with a normal performance for that degree of intelligence, and the relation of the two should be an indication of the obsessional characteristics only. This procedure is of course very different from attempting to measure several things at once without discriminating which has the greater weight in the final result. It is more akin to clinical procedure where, for instance, a battery of performance tests compared with a Binet test can indicate personal preferences, previous experience, or even *w* ("will") under certain circumstances.

Present Investigation

The present investigation provided an opportunity for observing unitary factors on the personality side in relation to the factors *g* and *p* ("perseveration") which were measured by Miss C. A. Simmons and Dr. W. Stephenson, and in relation to the marked differences of personality presented by patients who were under treatment at the Maudsley and other hospitals.

The questions to be studied were:

(1) What are the unitary or group factors to be assessed in considering personality?

(2) How may these factors be measured?

(3) How are they related to each other and to the total personality?

In the first place one unitary factor in the field of personality had already been discovered by the researches of Dr. Webb.[4] This factor was found by him to be common to all character qualities, particularly qualities such as perseverance, trustworthiness, and self-control. The letter *w* was used to designate this factor to denote its possible identification with "will." It had been assessed by means of estimates supplied by several observers; similar methods used in the present research verified the validity of the *w*-factor and revealed a second factor common to the following pairs of qualities:

elation-depression
confidence-diffidence
talkativeness-tendency to be silent
activity-inactivity
excitability-inexcitability.

These, among many other qualities, had been described in some detail and careful estimates supplied by doctors and nurses on graphic rating scales on which each of the above pairs represented opposing tendencies of the same trait. These qualities were more readily and consistently rated by the judges than were *w* qualities, different judges showing a considerable measure of agreement here, whereas on *w* qualities there was some disagreement. This new factor proved to be more easily measurable than *w*, and will be referred to here as *f* because the tests which we found to measure it depended largely upon the *flow* or *fluency* of the patients' responses.

The nature of the traits included in this group suggests at once that *f* is associated with extremes of mood and *f*-tests did in fact differentiate clearly manic and depressed patients where differences of mood are most marked. Though it would be premature at this stage to attempt to attach a verbal label to this factor, the importance of mood dominance in determining personality cannot be overlooked. Terms such as optimist and pessimist are frequently used in literature to denote personality types, and psychological writers constantly refer to such contrasted types. Professor C. Burt, for instance, points out two groups of emotional tendencies which he terms sthenic and asthenic,[5] and these terms would appear to fit the "fluency" factor admirably.

[4] E. Webb, "Character and Intelligence," *British Journal of Psychology, Monograph Supplements*, 1915, I, No. 3.
[5] C. Burt, *The Subnormal Mind* (Oxford University Press, 1935).

Recently, too, Dr. Emanuel Miller[6] has observed marked difference in "mobility" in children attending his clinic, i. e., those who "always approach the tests with superabundance of movement and restlessness, talking spontaneously and moving about in a way not suggestive of erratic and false movement, but rather of overflow." They are always in a state of tone. Others "always sat silently throughout the testing, almost toneless and with a minimum of movement."

One is even tempted to point out that the factors *w* and *f* would fit Professor James's classification of explosive and obstructed types with its two types of each:

the explosive type due to exaggerated impulsion (? high *f*)
the explosive type due to deficient inhibition (? low *w*)
the obstructed type due to deficient impulsion (? low *f*)
the obstructed type due to exaggerated inhibition (? high *w*)

These analogies are, however, somewhat speculative, and it is perhaps more profitable to examine the nature of the responses to tests which did actually agree with the judges' ratings on *f* qualities.

The *f* tests required associative rather than reasoned or logical responses and were scored for the number of associated items given by the patient in response to many different types of material such as words, pictures, ink blots. The tests were not all verbal, but a patient only made a high score when varied stimuli could touch off easily a quantity of associated material, and when he was able to indicate this material freely. The need for speed was stressed throughout, and the patients were urged to ignore quality. As Dr. Hubert[7] has pointed out, the number of items expressed did not always represent the number of items which occurred to the patient but only those he chose to express. This was well illustrated by one subject who, in a word-association test, gave the word "wall" and then paused for some time. When questioned about it later, she said that she had tried to find some word to express the texture of a particular wall surface, that several adjectives had occurred to her, but that she had rejected them because they did not convey the shade of meaning she had in mind. This is very typical of considerations which tended to reduce speed on the tests and give a low

[6] Emanuel Miller, "Temperamental Differences in Children," *British Journal of Educational Psychology*, 1933.

[7] W. H. de B. Hubert, "Some Psychiatric Considerations of the *f* Factor," *British Journal of Medical Psychology*, 1934, XIV, 131-136.

"fluency" score. Nevertheless, some degree of restraint of this kind seemed inevitable with normal subjects, and *f* seemed to measure less perhaps the restraining or inhibiting force than the superabundant surging pressure of mental activity which was able to overcome such barriers in the case of manic patients. Dr. Cattell[8] uses the term "surgency" to describe a similar group of traits which he examined in training college students.

It should be pointed out here that the *f* tests showed some correlation with *g* as measured by Miss Simmins, and the scores on the *f* tests were, for this reason, expressed in terms of the *relation* between crude scores on *f* tests and the *g* score. Neither *f* nor *w* factors as measured by estimates showed any relation to *g*, and the two personality factors *w* and *f* were proved to be entirely independent of each other.

Speed Tests

In order to throw some light on the processes involved in the *f* tests, other speed tests of a different nature had been applied to the same patients. These tests consisted of perfectly definite and easy tasks to be performed at maximum speed, e. g., reading easy words as rapidly as possible. Now, although these tests, like the *f* tests, depended upon the speed of the response, differences of behavior were very noticeable on the two types of test. Many depressed patients who made low scores on *f* tests worked extremely rapidly on these tests where there was no element of choice, or doubt, or quality, to be considered—in fact, some of the highest scores were made by depressed patients. When the scores on the two batteries (speed and *f*) were compared, some striking differences were seen for individual patients. On examination these differences proved of some interest. Cases with high scores for *f* ("fluency") but low scores for speed tended to be those patients who had been given low ratings for *w*-qualities (perseverance, industry, tenacity, etc.) and vice versa. This relationship was proved to be statistically significant and was in agreement with the fact, observed during the testing, that the easy speed tests depended largely upon the degree of effort made whereas on *f* tests effort was a hindrance rather than a help.

The relationship of the speed tests scores alone (i. e., not as compared with *f* tests) to traits of perseverance, industry, and de-

[8] R. B. Cattell, "Temperament Tests," *British Journal of Psychology*, 1933, XXIII, 308-329.

termination was very similar to that of scores on the *p* (perseveration) tests used by Dr. Stephenson. Both speed and *p* showed a small positive correlation barely significant perhaps, though suggestive; but when speed tests were scored in relation to *f* tests (*s-f*), the resulting correlation with *w* qualities was quite significant.

One further test which showed some correlation with *w* was a tapping test which demanded sustained effort after fatigue had set in. The patients were required to tap on a small lever as rapidly as possible and after several practice periods were urged to maintain this speed for two minutes. They were warned that it would be tiring, but were urged to maintain their speed at all costs. In almost every case the patient showed interest in his own score which was registered automatically and was visible throughout, and made some effort to maintain speed. This test is perhaps of some importance for future research on *w* in that the correlation for this single test with the perseverance ratings was as high as for the entire *p* (perseveration) battery, yet the test takes only five minutes to administer. When scores on the tapping test were added to the speed-versus-fluency scores referred to above, the correlation with *w* was considerably increased, again corroborating the fact that the same group of traits was being measured in each case.

This group, however, did not include the whole group of *w* qualities but only the three traits, perseverance, industry, and tenacity (or determination). Three other traits which have usually been included in the *w* group—self-control, trustworthiness, and unselfishness—showed no correlation with the tests, in spite of the fact that these traits as measured by estimates were very closely associated with the perseverance group.

The only score which gave a slight correlation with the self-control group of *w*-qualities was obtained from three tests scored for error alone without regard to output. It is well known that, for a given individual, increase of speed is usually accompanied by a corresponding increase in the number of errors. In comparing individuals it is therefore of some interest to note the errors alone, though of course few or many errors may indicate different factors at work in each case—a patient may work slowly because of extreme caution or because of lack of effort.

The three tests used for scoring errors included two where the patient was asked to work rapidly and a third where he was asked

to avoid error and allowed to work at his own rate. In all three tests excessive speed resulted in errors, and therefore the impulsive, hasty worker was at a disadvantage.

These three tests were:

(1) *Counting dots.*

Strips were shown with from twenty to thirty dots on each arranged in an irregular line as in Figure I.

FIGURE I

The patient was required to count them in ones as quickly as possible, pointing to each one in turn. (This device was adopted to minimize the influence of *g* because the more intelligent subjects tended to work rapidly by grouping the dots.)

(2) *Dotting on squared paper.*

The patient was asked to put a dot in each square as rapidly as possible, but not to miss squares or make dots on the lines.

(3) *A maze test.*

Here the patient was asked to move his pencil through the pathway shown in Figure II without touching the lines. No time limit was suggested, and the test was administered at the beginning of the test series before the speed tests had influenced the patients "set."

The total number of errors on all three tests was too small to give a normal distribution. Many patients made no errors at all and so were undifferentiated. In spite of this unsatisfactory feature, however, the correlation with ratings was not altogether insignificant. The correspondence with *w* qualities, though not high, was definitely suggestive, and the score was consistently related to *both* groups of *w* qualities—the perseverance group and the self-control group. In view of the need for measures of *w* which correlate with the total ratings, this score for errors seems promising, and if the tests could be modified to yield a larger number of errors without increasing the difficulty of the tests, such a score should be valuable in a battery of *w* tests.

Earlier in the investigation a large number of the existing techniques for measuring *w* had been tried out, but all showed insig-

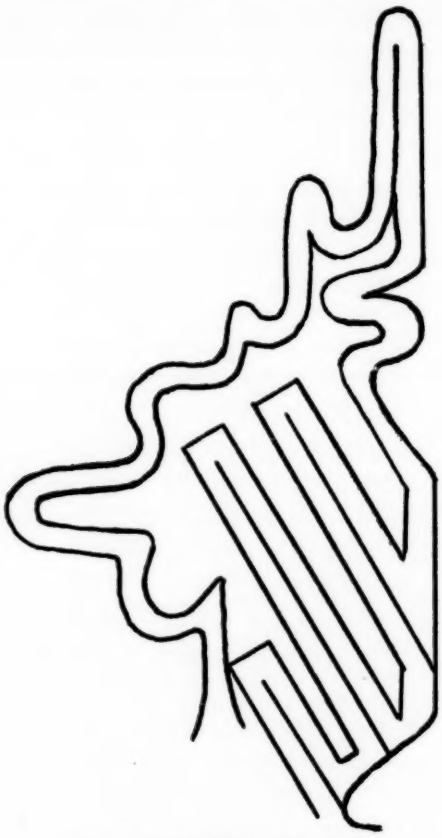

FIGURE II (Reduced in size)

nificant or negative correlations with estimates so that these tests for perseverance and the total *w* ratings seem worth following up.

CONCLUSION

Now that two factors on the personality side, *w* and *f*, have been established, and, to some extent, measures found, our knowledge of these factors can be increased by experimental procedures as in the case of *g*. As soon as it was proved that a general factor was measurable on the cognitive side, knowledge of the nature of that factor was obtainable. So in the present work, now that *f* appears to be measurable, the significance of mood changes in relation to personality can be studied. For instance, the normal ranges of elated and depressed reactions can be established, and individuals

may be tested from time to time to discover whether the rapidity of mood changes has any relationship to temperamental make-up or whether the range or rate of variation is characteristic of that individual.

Even in the case of *w* where the tests are less clearly proved, the data already available suggest the possibility that perseverance and industry depend to some degree upon *f* as well as *w*, whereas self-control, trustworthiness, and unselfishness may be regarded as purer *w* traits.

It is noticeable in this factor study that correspondence between tests and estimates has, in almost every case, been arrived at only when two or more factors have been compared. Thus *f* is only significant in relation to *g*, speed is only significant in relation to *f*, and *w* can only be satisfactorily assessed when *f* is taken into account. That this is true of individual measurements, as well as general tendencies, is illustrated by Dr. Stephenson's[9] analysis of a group of cases in respect of *p* and *f* where he finds a correspondence with clinical types when patients are examined in terms of the relations between *p* and *f*.[10]

[9] W. Stephenson, "Measurements of Mental Tension in Psychiatry," *British Journal of Medical Psychology*, 1934, XIV, 105-109.

[10] The correlations and other statistical material which support statements made in this paper may be found in "Studies in Experimental Psychiatry," *Journal of Menal Science*, 1935.

PERSEVERATION AND CHARACTER

W. STEPHENSON
Psychological Laboratory, University College, London

Many studies have been made in England on perseveration, and many more are still being made. Perseveration tests, which I prefer to call *p*-tests until I know for certain what the tests really measure, are by far the best tests of character, or measures of character, known to me. But I think that all is not well with this work, and in the present article I am to ask a few awkward questions about tests of perseveration, about their use as tests of character, and about the value of the concept of perseveration itself.

What Is Perseveration?

What I would regard as genuine perseveration is best seen in certain cases of mental illness, and the psychiatrist is precise about the symptoms so named. A perseveration, for him, consists of the persistent repetition of a word, or group of words, or of a motor action. This repetition occurs in spite of the patient's attempts to change it, and in spite of any attempts by the observer, doctor or nurse, to stay the perseveration by introducing new stimuli.[1]

A typical example would be that of a woman who, after eating her dinner, continued to manipulate her knife and fork as though food remained on her plate. I have observed a woman, after her bath, who continued for half an hour to make the motions of drying herself, after she had in fact been dried, clothed, and seated in her bed. Another, asked to write her name, scribbled it over and over again for nearly twenty minutes, with, apparently, the deepest interest and concentration.

After having ceased for a time, these perseverations may recur at a later time. Thus the patient last mentioned began to per-

[1] See Henderson and Gillespie, *A Textbook of Psychiatry*, Oxford Medical Publications (1927), p. 85.

severate the writing of her name again, two hours after it had ceased.

Perseverations are quite distinct from stereotypies. The latter consist of an action, or group of actions, or words monotonously repeated, or a posture maintained, long after fatigue would ordinarily have caused relaxation. The Egyptian mummy attitude of many dementia praecox patients is an example of a stereotypy. The observer can change or stop a stereotypy for the time being, whereas perseverations are not so alterable. Perseverations are certainly odd, but stereotypies are bizarre, fantastic, and ridiculous.

But the most striking difference between these two apparently similar kinds of activity has yet to be mentioned. It is one that textbooks of psychiatry are apt to overlook. The patient displaying a stereotypy has lost contact with reality, at least as far as the stereotypy is concerned, and usually completely. But the perseverating patient is still in touch with reality, is still subject to normal urges, impulses, and desires. Often, indeed, the only sign of this normal conative life that one can perceive in a dementia praecox patient, is seen when the patient is perseverating.

The patient I mentioned who perseverated the writing of her name illustrates this important difference. She perseverated almost every activity that she began. But those perseverated most, and more often recurred, which were matters of consequence to her. She was trying to please her doctors and nurses, and writing her name was charged excessively by this desire, because I had asked her to write her name. This volitional determination of a perseveration is psychologically its most illuminating feature.

Now perseverations are only rarely observed, even in a mental hospital with two thousand patients or more. There are at least a thousand other phenomena to be observed in such a hospital, from abstinence delirium on the one hand to rabid volubility on the other, from hysterical amnesia to feelings of unreality. Why, then, one may legitimately wonder, has this one rare symptom become the focus of so much attention in English psychological circles? During recent years at least forty studies on perseveration have appeared in England as papers or theses, and within my own knowledge at least twenty of our small band of active psychologists have conducted, or are engaged upon, studies on this topic. What is the source of its interest for so many workers?

What Is Character?

The answer to the above queries lies in the suggestion that perseveration is somehow connected with character.

We distinguish nowadays between character and temperament. Character is estimated as Spearman's, or Webb's, *w*-factor,[2] whereas temperament is estimated as Garnett's *c*-factor,[3] and measured to some extent objectively by tests of fluency.[4] Character is concerned with volition; temperament, rather with affects and emotions. Character deals with one's motives, especially those newly determined by a will act and issuing in a determining tendency, in a striving to attain ends, and in the keeping of goals before one's mind—whether the goals are idealistic and moral, or frankly otherwise.

Character is estimated in terms of traits such as persistence of motives, perseverance in the face of difficulties, trustworthiness, conscientiousness, tendency not to abandon tasks from mere changeability. Temperament, on the other hand, is estimated in terms of cheerfulness, good humor, and optimism, sanguineness, tendency to anger, temper, physical energy, and, at the other end of the scale, in terms of dull phlegmatism and melancholy. Temperament indeed reads nearly the same as the old classical temperaments, in a linear scale from choleric to melancholy.

The Facts

The facts are certain enough that *p*-tests, called tests of perseveration by most workers (by which they must mean that these *p*-tests measure perseveration), correlate with estimates of character. I shall discuss these *p*-tests later, to see whether they really do measure perseveration.

It is interesting, on the other hand, that tests of *f*-factor (fluency) correlate positively with estimates of temperament, and not at all with estimates of character. *P*-tests, however, frequently correlate with estimates of temperament, as well as with estimates of character.

If a group of individuals is estimated for character qualities,

[2] E. Webb, "Character and Intelligence," *British Journal of Psychology, Monograph Supplements*, 1915, I, No. 3.

[3] M. Garnett, "General Ability, Cleverness and Purpose," *British Journal of Psychology*, 1919, IX, 354.

[4] G. L. Studman, *Journal of Mental Science*, 1935, LXXXI, 107; and the preceding article in this journal by the same author.

these correlate negatively with the scores supplied by *p*-tests of the "Triangles" and "Writing S2" kind, described by Bernstein[5] and the present writer.[6] The correlation, however, is often not linear. The main trend is certainly negative, but extremes of *p*-score are associated with poor character while medium *p*-scores are associated with excellency of character.[7]

Now I am not much concerned about this startling non-linear correlation. Even Cattell[8] has just decided that the extremes of *p*-score are associated with the same kind of poor character because at one moment these individuals are high perseverators, while at the next they are the very opposite, low perseverators. This is in line with an earlier finding of my own, that individuals displaying perseveration in the most vivid way in *p*-tests, markedly fluctuate in just such a manner, from extremely high to low (or better, to *no*) perseveration.[9] Moreover, *p*-scores and *f*-scores (for fluency tests) not infrequently correlate, and it can be shown that this accounts for some of the non-linearity between *p*-scores and character estimates.

The main fact to which I wish to draw attention is merely that *p*-score is certainly associated with character, to the extent that high *p*-score goes with poor character. Indeed, recent work shows that all the poorer qualities, poor intelligence, poor self-control, poor powers of concentration, poor physical energy, weak character in general, are associated with high *p*-score.[10]

Now high *p*-score means, apparently, marked perseveration. It is the greatest perseverator who has the weakest character.

But it is just here that an assumption is being made. Do *p*-tests indeed measure perseveration? Are the *p*-scores really scores of perseveration? And even if we can answer this question in the affirmative, care is needed when we consider the implications of the correlation between character and perseveration. It would be tempting, and all too easy, to assume that an individual's character is weak because of his high perseveration, the latter being a cause to some extent of character deficiency. The truth may be, however, that an individual

[5] E. Bernstein, "Quickness and Intelligence," *British Journal of Psychology, Monograph Supplements,* III, 7, 1924.

[6] W. Stephenson, "An Introduction to So-called Motor Tests of Perseveration," *British Journal of Psychology,* 1934, IV, 186.

[7] R. B. Cattell, *Character and Personality,* 1934, III, 154-163.

[8] R. B. Cattell, *Journal of Mental Science,* 1935.

[9] W. Stephenson, "*P*-score and Inhibition for High-*p* Praecox Cases," *Journal of Mental Science,* 1932, LXXVIII, 908.

[10] A. Maginess, *A Comparison of Orectic Factors for Normal and Mentally Defective Children.* Thesis (University London Library, 1935).

displays perseveration only because of his weak character, and that character owes no more to perseveration than it owes to the man in the moon.

Perseveration Observed in *P*-tests

First, then, let us answer the question: do *p*-tests measure perseveration?

Let anyone try to write the following upright and inverted triangles, one after the other, at very great speed:

△ ▽ △ ▽ △ ▽

It frequently occurs, especially if we are fatigued, that we find it impossible to do the required alternation from one triangle to the other. For a few moments we repeatedly write only one kind of triangle, and find it impossible to write its opposite. We know that we are not doing the required alternating, but for the moment we simply cannot break away from writing only the one kind. The more we strive to break away from this state, the less likely are we to succeed. In short, for the moment we are perseverating. Often, to stop ourselves from this perseverating, we have to stop writing altogether, and then begin again with a new mental set.

This peculiar state must not be confused, as it often is, with the mere writing of one kind of triangle because we have forgotten that alternations are required. Incidentally, it is easy to see that the individual of weak character is apt to lapse in this very way, because of the weak mental set that he has made at the outset.

Extraordinary instances of this motor perseveration, as it is called, can be seen if certain dementia praecox patients can be got to attempt a *p*-test like the "Triangles" test. I have watched a patient trying to write an inverted triangle, who showed clearly that she knew what was required, who could write the upright triangle with ease, who repeatedly made the necessary preliminary steps towards writing the inverted one, and yet who failed time and time again. The attempt to write an inverted triangle resulted over and over in the writing of an upright one, in spite of every effort to do otherwise. The perseveration was often accompanied by mutterings of exasperation, and almost of despair, showing that there was no lack of desire to succeed.[11]

But these instances of profound perseveration are by no means common, and they are sporadic only. At one moment that patient

[11] W. Stephenson, *Journal of Mental Science, loc. cit.*

may be in the depths of perseveration, while in the next she may be completely free from it. It is the same with normal individuals. The instances of perseveration are short-lived lapses on an otherwise unruffled surface of successful alternating or switching activity.

Thus, instances of perseveration certainly occur in *p*-tests of the "Triangles," "Writing S2," and "Writing ZZ" kind. Yet there is evidence to show that the *p*-scores in the main are not measures of this perseveration, whatever else they measure, evidence that I am to publish elsewhere. Moreover, even if they were, the perseveration is merely incidental to something of much greater interest for the study of character. It can be shown that character determines the *p*-score, and that tests other than presumed tests of perseveration serve just as well as *p*-tests to give correlation with estimates of character. We can show that the difficulty experienced when switching in a *p* test is dependent upon the effort expended upon it. He frequently finds it most difficult who tries hardest to succeed, and this explains why many individuals of excellent character obtain a medium *p*-score. He often succeeds best who scarcely tries at all. This explains why many individuals of poor character obtain low *p*-scores.[12]

Explaining the Correlation between P-Scores and Character Estimates

Otto Gross, and many who have followed in his footsteps during the present century, believed that the high perseverator (or the individual with high *secondary* function, as he called it, a subject to which I revert later) should have a well-knit character. As has been shown, if *p*-tests really measured perseveration, then the very opposite would be true, that the person with highest perseveration has the poorest character.

Lankes,[13] however, far from considering that perseveration contributed to the shaping of character, concluded that character merely moulded one's perseveration. The individual of firm character, according to Lankes, could modify and control his innate tendency to high perseveration. Presumably we are all high perseverators. But only the individual of weak character retains this high perseveration unalloyed, while the individual of strong character modifies his.

Until recently, however, workers have not had Lankes's caution.

[12] Low *p*-scores are obtained by him who succeeds best in a *p*-test; high *p*-scores, by him who is least successful.

[13] W. Lankes, "Perseveration," *British Journal of Psychology*, 1915, VII, 387.

It has been implied, in the writings of Pinard, Cattell, and others, that perseveration is a cause, a moulder, a former of character. Perseveration is not synonymous with perseverance. Instead, it is its very opposite. Perseveration destroys any hope of perseverance and good character. It is epitomized by dull brooding, by goal fixation, and by narrowed motives. Only now, in his most recent paper, Cattell has concluded otherwise, that the individual's perseveration is a consequence of his experiences.[14]

Now many tests which do not involve perseveration at all, on any theory of perseveration, inertia, secondary function, or whatnot, serve just as well as *p*-tests to give negative correlation with estimates of character. If we measure individuals for easy speeded activity (such as ability to cancel all the letters on a page of print), and also for fluency (such as giving words beginning with L), the difference between the two can supply correlation with *p*-tests, and negative correlation with *w*-estimates. Again, as Lankes found, but did not draw the obvious conclusion, and as Cattell and myself have since repeated, certain cancellation *p*-tests act only in one direction. One first cancels XOI's from a page of mixed letters; having practiced this, one next cancels HGVN's from the same page. It is supposed that the former cancelling tends to disturb the latter; the high preseverator will tend to cancel XOI's when he should be cancelling HGVN's, due to the persistence or perseveration of the former. Actually one can cancel HGVN's first, without ever cancelling XOI's, and the test is still exactly as it was before—it correlates with other *p*-tests and with *w*-estimates, in spite of the fact that perseveration could not have entered into it.

The truth is that *p*-tests, first and foremost, present miniature life-situations to the individuals who attempt them. They offer a slight difficulty to be overcome, and allow for the interplay of conation and motivation, of effort, purpose, will, in a word of character. *P*-tests are thus direct tests of character, and it is not necessary to consider that perseveration plays any essential part in them.

Perseverative Tendency and Determining Tendency

Perseveration, then, seems to be merely the minor symptom I first described. All its presumed contact with character would seem to be a fiction. But psychologists as great as Gross, Wiersma, Jung, and Spearman, have used the concept of perseveration, or a derivative of it, as an essential, if not all-important feature of human

[14] R. B. Cattell, *Journal of Mental Science, loc. cit.*

behavior. If perseveration is a minor symptom only, what is it, then, that has been mistaken for it in the conduct of human beings?

The question can be answered by referring to two concrete examples. Two children are playing before me, one my daughter, aged 27 months, and the other her cousin, a boy of the same age. Both are fit, lively, intelligent children. But they are strikingly different in temperament and character. The girl is quiet, reserved, deliberate in her actions. She is shy with strangers, and has developed a mass of subtle emotional responses and nuances much in advance of other girls of her age. She plays coherently, there being little of irresponsible shifting from one focus of attention to another. She dwells repeatedly on an idea that takes her interest, and her desires are deep and long lasting. But she is apt to brood under stress, is jealous of her cousin, and has shown him no friendliness during his fortnight's visit to our home.

The boy is of a very different stamp. Aggression, free outbursts of anger, elementary expression of his emotions with none of the subtlety of the girl, and yet withall a noisy cheerfulness and good humor and friendliness that reminds one of a diminutive Falstaff—this is the boy. He plays few games for more than a few moments, his interest vacillates, he knows little of shyness and jealousy.

The ready smile and beaming cheerfulness of the boy, his abundant energy and extroversion in general, clearly indicate a temperament very different from that of the girl. I would estimate him high for temperament (*c*-factor). The girl, on the other hand, shows much more coherency of self-willed action, more persistence of her own freely willed motives and desires. The boy's immediate interests, motives, desires, and instincts are no less strong than the girls, for in a way they are stronger and more insistent. But the girl has made a cognitive and affective fine art of hers; each seems to outlast its immediate stimulation, each lasts longer, each recurs, each seems to be woven intimately into her personality.

It is this latter quality of behavior that Gross was describing when he talked of secondary function; that Müller and Pilzecker considered to be determined by perseverative tendency; and that Lankes might have explained as the result of the individual's tendency to high perseveration in general. All these, in keeping with the times, were psycho-physical explanations of the coherency and persistence just described. One had only to show, finally, that individuals characterized by this kind of behavior were also high per-

severators, and the theory could be held that perseveration is the underlying cause of this quality of behavior.

The concept of perseveration again, however, or any of the others just mentioned, fails to do justice to the facts. These concepts neglect in their description the essential volitional features entailed in this kind of coherent conduct; it is only those ideas, or activities, charged with desire and wishes that show this persistence. What we see is a feature of the child's volitional life, of its urges, desires, and impulses. The word *perseveration,* or perseverative tendency, in no way draws attention to this all-important fact. Nor do these concepts help us to understand the behavior under examination.

I would link all this persistence quality to the concept of determining tendency. Ach, whose work I greatly respect, has concluded that perseveration and determining tendency are not one and the same thing, and with this I am in agreement. Genuine perseveration in no way helps me to understand character, nor, as I have suggested, does it enter into the tests that measure character. But the concept of determining tendency does help me. The child, from its own volition, sets up a state of tension, which normally issues in striving and conation with the end that a desired goal be reached, or satisfaction acquired, and the determining tendency or mental set tends only to end when the goal is reached. The concept at once brings us into the sphere of volition and conation, and it is precisely in this sphere that we hope to find the secrets of character.

The difference between the two children I would explain, then, in terms of determining tendency and its properties, and not in terms of secondary function, perseverative tendency, perseveration, and the like. It remains only to ask what is the connection between determining tendency and character. But this, like the serial story in a newspaper, must be continued in another article.

Finally, lest anyone should think otherwise, my criticisms in no way lessen the respect due to *p*-tests. Indeed they should only heighten it, since still more clearly than before do we see their connection with character. Moreover, the very success of *p*-tests contradicts the pessimistic artistry of those psychologists, like Allport and Vernon,[15] who have decided that artificial tests cannot touch the sacred portals of temperament and character. For *p*-tests are artificial in all conscience, but they certainly correlate with character.

[15] Allport and Vernon, *Psychological Bulletin,* 1930, XXVII, 677; and P. E. Vernon, *British Journal of Psychology,* 1933, XXIV, 156.

RACE AND MODE OF EXPRESSION:
A PRELIMINARY INVESTIGATION IN COLLECTIVE PERSONALITY

A. A. ROBACK
Cambridge, Massachusetts

Ever since the branch of collective psychology was formed, the problem whether there is a characteristic mode of expression in members of a given race has attracted the attention of thinkers along collective lines. Where the individuals involved reside in territory where their race forms a compact group, the problem is easy of solution, as evidenced by the stupendous work on Folk Psychology by Wundt alone, not to mention the vast literature on race productions which was inaugurated in Lazarus and Steinthal's *Zeitschrift für Sprachwissenschaft und Völkerpsychologie.*

The question assumes a characteristically complicated aspect when we are dealing with individuals of a certain race who have become expatriate or, better, have been born into a foreign milieu. The Jews offer probably the best illustration of this class. As they become more emancipated, they become also more assimilated; and even though they may express their sympathies with their race or even identify themselves with its interests, nevertheless they absorb more and more the qualities of their milieu. Not the most ardent devotee of Jewish nationalism will deny this fact.

The crux of the matter, however, is not reached until we ask whether amidst this assimilation there may not be a residue which resists all efforts at amalgamation and which must therefore be recognized as the permanent core of national expression. When pressed for a reason, the theoretical nationalist would fain ground the differentiation in the different unit characters of the germplasm. To demonstrate this point biologically would be possible perhaps only after unit characters could be studied under the microscope, but meanwhile the controversy is still surrounded by speculation. On

the one hand, there are the mechanists and materialists who stake all their wares on the potency of the environment. On the other hand, the mentalists and their allies, while prepared to concede that the environment does play an important part in shaping one's style of thinking, are nevertheless committed to the view that no individual can escape the racial *élan* which, like the Gulf Stream, is not swallowed up by the bodies through which it flows.

The question about the purity of the Jewish race tends to complicate the issue all the more, and it goes without saying, of course, that an individual who is only one-eighth or one-sixteenth Jewish cannot, by any canon of reasoning, be regarded as a representative of the Jews. Even, however, if we concede a certain amount of miscegenation in past centuries, it would seem that the general question stands with regard to a particular outstanding individual of the race, whose ancestry, so far as we know, has been Jewish.

Methods

One of the first to discuss the problem with regard to Jewish specificity in cultural productions has been Zhitlowsky in his lengthy essay, "Yid un Mentsh," which appeared about a quarter of a century ago. In this address, as it originally was, he set himself to the task of discovering whether Jewish thinkers in various countries and ages reveal a common bond in their approach to a problem, even if not in their ideology, and he answers the question in the affirmative. He finds a tendency in Ibn Gabirol, Spinoza, Marx, and Bergson to separate the basic concept of their field into two subsidiary phases, which, after considerable analysis and elaboration, they attempt to synthesize into a monistic whole. The illustrations, it must be said, are scant and unconvincing, yet at least the thesis had been laid down.

What is perhaps more suggestive is the reference to Möbius as placing the Jewish mathematicians in a category with the French, who are said to be analytic and favoring algebra, while the German mathematicians are regarded as intuitive and excelling in geometry.

I was not able to verify the reference in Möbius's *Anlage zur Mathematik,* although several years ago, wishing to check up on Zhitlowsky's allusion, I did come across an *alleged* categorization by Möbius, but was unsuccessful in tracing this passage in my recent reading of the same book. With regard to the Jews Möbius maintains a dubious reserve even to the extent of saying that "it has

repeatedly been pointed out that among the Jews there were a good many excellent mathematicians, yet it is difficult to draw any conclusion from this" (*Anlage zur Mathematik,* pages 111-112).

It has occurred to me that we should never get very far in the solution of our problem if we keep generalizing or making inferences *ex post facto.* Theoretically it stands to reason, as I have argued at the Oxford Congress of Philosophy in a paper entitled "When Has Jewish Philosophy Become Extinct?" that the Jews cannot with justice be made an exception when the history of philosophy is conceived in national perspectives, and that the Jewish thinker can no more escape from the germplastic impositions of his race than the German and the Frenchman can free themselves from their respective cast of thought. It is true that the noted Spinoza scholar, the late Professor Gebhardt, took exception on the ground that the Jews were not a unitary people but were made up of two distinct groups, viz., the *Ashkenazim* and the *Sephardim.* Nevertheless, the fact that the Jews are recognized as Jews on less impressive occasions should, as I retorted then, render them just as homogeneous when it is a matter of receiving credit for a doctrine, law, or principle. In other words, the application of the saying, "What is sauce for the gander," etc, should be seconded by the further observation that we cannot be fish at one time and fowl at another.

Paradoxical, though it be, most of the major Jewish philosophers of the present day are not willing to own up to any Jewish influence in their philosophy. Bergson, who has often been the butt of his opponents' anti-Jewish bias, in an interview some twenty years ago denied that there was anything Semitic either in his ideas or in his style.[1] Lévy-Bruhl similarly wrote me that he thought his philosophy was distinctly French, and the late Émile Meyerson was equally emphatic in his affirmation that the tradition of the milieu was instrumental in the shaping of his system. In reply to all my pleading, he would demand the proof. But to offer proof would require a searching investigation; for naturally the impetus of the national

[1] "I do not much believe in permanent special qualities of races. Nature is very often nothing else than habit and education. There are racial differences between the white, yellow and black races, but there is no difference in the white races. People can adopt the qualities, the defects and the habits of the people among whom they live. In Europe we see that the difference in races is nothing but habit, education and the degree of living together.

"I doubt whether the Jews have any special hereditary defects or qualities, considering that their blood has been so mixed—very much more mixed than is believed. Whole tribes in Russia were converted to Judaism" (From an interview with Henri Bergson as reported by Herman Bernstein in 1912).

tradition does necessarily outweigh the subtle influence of the race; and much, in any case, would depend on interpretation. Under the circumstances, the controversy could be drawn out to interminable length without yielding any fruitful, let alone, decisive results.

So far we have been dealing with two methods to establish the racial specificity in cultural productions of individuals creating in a non-Jewish milieu:

(1) The *deductive,* emanating from the conception of the term "race."

(2) The *analytic,* i.e., the analysis and comparison of not only Jewish thought in various countries but in addition the works of Jewish and non-Jewish thinkers in a given country.

It is evident that the first method, of the a priori variety, will not carry much weight with those who are altogether skeptical of the inherent qualities and hereditary power of race in general. The second method, while more satisfactory, can scarcely offer conclusive evidence until the investigation is undertaken on a gigantic scale; and even then, doubting Thomases will see coincidence here and individual quirks there, without granting the claim.

A third method, however, may be resorted to which, if conducted vigorously in various fields and over a sufficient period of time, would be bound to decide the issue. This method may be called the experimental or pragmatic approach.

Thus far I have not been able to pursue this method except in a preliminary way, but I may at least describe it for the benefit of those who should desire to try it out under favorable conditions.

For some time I had been thinking that the real test of the issue at hand lay in pointing out Jewish characteristics in material whose authors are unknown to the investigator. In other words, *racial detection or diagnosis in a sufficiently large number of instances where the elements are varied should clinch the point.* Naturally the content of the material might sometimes serve as a clue to the nationality of the writer, and for this reason we ought to experiment with matter-of-fact material. What is there more colorless than an examination paper where the student merely disgorges, so to speak, what he has ingested, and in some cases only badly digested, from the lectures and textbook? Here, if anywhere, we should have an objective basis for our inquiry; and if the Jewish examination books, shorn of all identifying marks, could be spotted from among

the rest, then surely creative productions should stand greater chance of racial tagging and therefore should be an indication of racial influence.

My colleague, Professor G. W. Allport, was kind enough to offer the examination books of his class in the psychology of personality both at Harvard University and Radcliffe College for this investigation. In all there were 64 examination books. After numbering each book—in many cases a double or triple book—Dr. Allport had the cover which bore both the name and the grade of the student torn off, so that there remained no external clue of the writer. About 20% of the students were Jews, but this proportion was not divulged to the investigator. It must also be considered that the Jewish students at Harvard University and Radcliffe College, particularly the upper class men and women, because of a rigid system of selection, one might say almost natural selection) are not given to the use of foreign idioms or barbarisms, so that the language factor would not enter here to betray the Jewishness of the writers. Every endeavor was made as far as possible to rule out any variable which might invalidate the results or at least call them into question. Professor Allport had thought that the reputed excellence of Jewish examination books might serve as a clue, but in the first place the mental set of the investigator was, and naturally would be in all such cases, to disregard the quality of the answers, and secondly, it transpired afterwards that some of the Jewish grades were C's, while many of the non-Jewish grades were A's and B's, so that it would have been risky to draw any conclusions upon such an indirect, if not flimsy, basis.

Results

The results, although not conclusive, are significant. Of the 64 examination books, all but 21 had been correctly matched as either of Jewish or non-Jewish authorship. As Dr. Allport, who checked the markings on the books with his key to which only he had access, observed, the correlation is certainly much above chance.

Dr. Rulon, of the School of Education of Harvard University, was kind enough to work out the probability relation involved in these results and reported that the probability was 3 to 1000 that the 43 correct markings could have been the result of chance.

Naturally one test of this sort, even were there 200 books instead of 64, could not be decisive. The reliability of the results depends a

good deal on a uniformly significant correlation. The percentage of correct answers does not necessarily have to be uniformly high, but relatively so. Hence another set of examination books, in a social psychology course, this time evenly divided as to Jewish and non-Jewish stock and scholastic grade, was presented to the investigator for the purpose of matching. It must be said that the results in this case were negative, scarcely any better than one might expect on a hit-and-miss basis. The difference in the two sets of books will be touched upon later. As the second step in the experiment involved a variation of the first step, the negative results are of course telling, but only more trials introducing like variations can determine the question of racial influence in cultural productions of assimilated individuals.

The Specific Signs

What were the indications in the examination books which would enable a reader to determine their Jewish authorship? Certainly we cannot leave it to such meaningless phrases as "a trained sense," "a certain undefinable something," or appeal to that factotum of philosophy and psychology, "intuition." Unless we have definite criteria which can be applied by trained observers so as to yield uniform results, there is no solution.

At this stage of the experiment, strange as this admission may seem, I cannot state what constitutes the earmarks of the Jewish mode of expression. I can merely offer what I noted in my approach, but it goes without saying that in this particular type of investigation, which in a sense is pioneer work, I am but a novice. I started out with a certain "feeling" or "hunch" of what to look for; and the number of errors in judgment of racial stock might well point to the fact that my "feeling" or "orientation" was not altogether well founded. The hypothesis, in other words, will have to be modified in the light of subsequent research.

To begin with:

(1) I noticed that where there was an option on the examination paper, some of the students (which I took to be Jewish) chose the broader, more historical questions instead of the technical and restricted questions.

(2) The Jewish students, I thought, were more expansive, not that they were more verbose, but introduced illustrations from their readings and experiences. The non-Jewish students seemed to be more matter-of-fact.

(3) The Jewish student appeared to show greater flexibility or elasticity in style.

(4) There is greater subjectivity in the replies of the Jewish students.

I might have been able to allude to differences in handwriting, if the subject were not so complicated. The impression I received was that the Jewish students write a more uniform, simplified, and legible hand than the non-Jewish students, or, to put it somewhat differently, there is a proportionally greater deviation from the normal copybook script in the writing of non-Jewish students. In one or two cases, where there was a doubt in my mind I decided that the backhand style or unusually large characters were not characteristic of Jewish writing.

The low correlation in the results of the second set of examination books must be explained on the basis of the more rigid and technical questions set. The subject offered less latitude. We must also remember that there were only twelve books, six of Jewish and six of non-Jewish students.

Suggestions and Recommendations

The next step in the training of the observer should have been the study of those examination books which were incorrectly classified. Only by such comparison can we hope to discover the specific marks of Jewishness and non-Jewishness, and unless we have made repeated trials with negative results, we are not justified in denying that there are such marks.

As already intimated, even if the results had been negative with examination books for material, the issue would not have been settled, for the reason that examination answers tend to reflect the class lectures or the textbook.

We may well devise a controlled method for classifying short stories, poetry, essays, *feuilletons,* music, etc., according to racial stock. Such creative productions would have to be of course relatively new, at any rate, unknown to the observers.

In another instance, investigations may be planned in schools where there is a large foreign element, using compositions for material. It is deplorable that there are no Jewish foundations that might be appealed to for subsidizing such research.

In conclusion let me explain, lest there arise some misunderstanding, that my views on probable racial differentiation in individual creations are not restricted to Jews and Gentiles. Such differences should be looked into wherever there are individuals of various stocks creating in a common milieu. Naturally some stocks will suggest a greater disparity than others, and the Jewish-Gentile dichotomy is likely to stay with us culturally.

TWIN-SIMILARITIES IN EMOTIONAL TRAITS*

HAROLD D. CARTER
Institute of Child Welfare University of California

The fact that identical twins are very similar in intellectual traits, as well as in physique, while dissimilar twins may be as different in intellectual abilities as in physical appearance, has been established in the literature. But the literature is lacking in adequate studies concerning the emotional attitudes, as contrasted with abilities. Such study is important, because information is desired concerning the effects of training and of hereditary predisposition in determining emotional attitudes. Studies of twins are desirable first steps in this field of research, since identical twins are always approximately the same in hereditary characters, and dissimilar twins may be greatly different in genetic make-up. When the basic facts about twin-similarities in emotional attitudes are known, it may be possible to infer the influence of heredity upon development of these traits. Our present purpose is to take a step in this direction by investigating twin-similarities in those feelings and emotions which are important in many phases of everyday life.

The traits selected for study are attitudes such as self-confidence, sociability, introversion, dominance, and self-sufficiency. A thorough study of twin-similarities in such attitudes requires at least an attempt to answer several questions. First, how do twins compare with non-twin populations in these traits? Perhaps there is a tendency for both twins either to be emotionally well-adjusted, or maladjusted. Second, how similar are identical twins and fraternal twins in these emotionalized attitudes? This question deals with the relative effects of heredity in determining emotional attitudes. Third, are the essential similarities of twins to be found in general attitudes and broad aspects of behavior? Or, instead, are the sim-

* The data for this study were collected by the writer in 1930-1931, while he was engaged in studies of twins, as Social Science Research Council Fellow, at Stanford University.

ilarities essentially those revealed in rather isolated and independent bits of behavior?

The presence of twins in a family serves to create a special social situation. Further, to one who has done intensive field work with twins, it is obvious that the social factors involved are not the same for the identical and fraternal twin pairs. The general similarities and dissimilarities of twins (for example, in physique) which obtain in particular cases are important factors in leading them to regard each other as equals, leader and follower, superior and inferior, friendly rivals, or unfriendly rivals. These considerations make it desirable to study twins with respect to emotional attitudes involved in social situations.

Difficulty in Securing the Information

Increase of knowledge in this field must depend upon improvement in methods of securing the necessary information. In studies of emotional or temperamental traits in twins, there has been much dependence upon questionnaires and reports of various kinds. For example, the answers of mothers of twins to the question whether the twins are alike or different in various traits have furnished the basis for some investigations. However, since there are emotional attitudes, such as feelings of insecurity, which can be observed accurately only by the person experiencing them, in some studies the twins have been asked to furnish self-reports. Errors inherent in such reports make it necessary to use standardized procedures developed for the purpose of minimizing the errors. The techniques of asking questions, and of treating results, have been improved in many ways. Despite admitted shortcomings, many interesting and important facts have been discovered.

For example, if an individual is asked bluntly whether he is happy or unhappy, the answer will often be an attempt to cover up real feelings. It is commonly true that blunt questions will not elicit correct information about emotional reactions. This fact is generally known; hence in the attempt to secure more reliable data, the makers of inventories have inserted many questions of a milder sort in a long list of questions bearing upon the same attitude. These are mixed with other questions which bear upon entirely different topics. By these methods, the individual tested is induced to give more exact information, because emotionally charged questions of too obvious import are not the sole content of the inventory. In-

ventories using these devices have many questions bearing upon each attitude studied; this makes it possible to secure quantitative results, through summation of scores for questions relevant to the attitude in mind. The Bernreuter Personality Inventory, used in the present investigation of the emotional attitudes of twins, is an example of this type of questionnaire.

The Bernreuter Personality Inventory

The Bernreuter Inventory consists of a list of 125 questions to be answered in one of three ways, by circling *Yes, No,* or *?*. Examples of the inventory questions are given in Tables V and VI of this report.[1] The questionnaire yields information about such traits as dominance, introversion, neurotic tendencies, self-sufficiency, confidence, and sociability, which are measured by the specialized scoring schemes.

A person's score for any one trait, such as dominance, depends upon his answers to many different questions. If an individual reports that he objects when a person steps in front of him in a line of people, the response indicates dominance and receives a positive score. Likewise, if he reports that he tries to get his own way even if he has to fight for it, dominance is again indicated. But if he is troubled with shyness, a lack of dominance is indicated, and a negative score is given. Since a large number of questions furnishes some indication of dominance, an individual's score for this trait is obtained by summing all the scores on the relevant questions. This technique results in a total score which is intended to represent the general level of dominance characteristic of the individual. It is of course possible that persons dominant in one situation will be submissive in others; hence the total score may be based upon some positive and some negative scores; to the extent that this is true, the total score is deceptive. These considerations make it desirable to analyze the responses to individual questions, as well as the total scores.

A fundamental assumption underlying the use of this test is that the behavior of an individual in a given situation is a useful indication of more than one trait. For example, an individual who reports that he usually avoids asking advice may indicate thereby that he is unsocial, and also that he is dominant and self-confident. Likewise, if an individual reports that books have been more entertaining than

[1] All tables are presented at the end of the article.

companions, the report may indicate that he is introverted, self-sufficient, and unsocial. Such tentative indications become more and more reliable indications of fact when supported by similar information obtained from increasing numbers of other questions. That is, the suggestions about traits, derived from answers to single questions, are checked by information from answers to other questions which approach the same traits from other situations. Each question contributes something to the measurement of several personality traits which are components in the total overt response.

Nature of the Traits

Several scoring schemes[2] have been developed for use with this test, the items receiving in each case the weights which are proportional to their relative value as indications of the several traits. Questions which have no bearing upon a particular trait are given a zero weight on the scoring key; weights for items have been determined through study of selected groups of individuals. For example, weights for the dominance scale are determined through study of a group of very dominant persons, in comparison with a group of very submissive persons.

The first test of this fourfold inventory is described by the author[3] as a measure of neurotic tendency. He states that persons scoring high on this scale tend to be emotionally unstable, and those scoring low tend to be very well balanced emotionally. Persons who report that they often feel just miserable, that their feelings are easily hurt, that they worry too long over humiliating experiences, that they are troubled by feelings of inferiority, etc., are likely to receive high scores on the scale for neurotic tendencies.

The second test is described as a measure of self-sufficiency. Bernreuter states that persons scoring high on this scale prefer to be alone, rarely ask for sympathy or encouragement, and tend to ignore the advice of others. Those scoring low dislike solitude and often seek advice and encouragement. Item responses which are heavily weighted on the self-sufficiency scale are those indicating liking for evenings alone, preference for solitary work, tendency to get ideas through reading rather than through conversation, tendency to face troubles alone, etc.

The third test is a measure of introversion. Persons scoring

[2] R. G. Bernreuter, *Manual for the Personality Inventory* (Stanford University, Calif., Stanford University Press, 1935).
[3] *Ibid.*

high on this scale tend to be imaginative and tend to live within themselves. Those scoring low are extroverted; they rarely worry, seldom suffer emotional upsets, and rarely substitute day-dreaming for action. Scores on the introversion scale tend to be high for persons who report that they daydream frequently, that they cannot stand criticism without feeling hurt, that they blush often, that they worry over possible misfortunes, etc.

The fourth test is one of dominance (self-assertiveness). Persons scoring high tend to dominate others in face-to-face situations; those scoring low tend to be submissive. Responses of the sort producing high scores on the dominance scale are those indicating lack of shyness, ease in conversing with strangers, tendency to take the lead in social situations, lack of inferiority attitudes, etc.

Interrelationships of the Traits

The above-described measures have been validated against previously existing tests designed to measure the traits named; the measures are not independent of each other. One might infer this from the description. Dominance and self-sufficiency as defined are positively but not highly correlated, while each has some negative correlation with the two remaining traits. The test of neurotic tendencies and that of introversion agree so well that they may be regarded as two slightly different measures of the same thing.

Two new scales developed by Flanagan[4] for use with this test have but recently been published.[5] In using these new scales, one analyzes the same responses to questions, but from a new point of view, in terms of two mutually independent personality traits. The first is described as a measure of confidence in oneself. Persons scoring high on this scale tend to be too self-consicous and to have feelings of inferiority; those scoring low tend to be wholesomely self-confident and to be very well adjusted to their environment. The second measure is one of sociability. Persons scoring high on this scale tend to be non-social, solitary, or independent; those scoring low tend to be sociable and gregarious. These two scoring schemes have been developed by application of Hotelling's method of principal components.[6] Flanagan's two traits are independent of each other, but not independent of the four Bernreuter scales.

[4] J. C. Flanagan, *Factor Analysis in the Study of Personality* (Stanford University, Calif., Stanford University Press, 1935).

[5] R. G. Bernreuter, *Manual for the Personality Inventory.*

[6] H. Hotelling, "Analysis of a Complex of Statistical Variables into Principal Components," *Journal of Educational Psychology,* 1933, XXIV, 417-441.

The intercorrelations of the Bernreuter scales have been published elsewhere,[7] and for the particular sampling used in the present study these intercorrelations are given in Table III. It will be apparent to the reader that in using the six different scoring schemes one is really regarding all the results (i.e., all the answers to all the questions) from the point of view of each of six everyday concepts. The test may be regarded as a standardized sampling of self-report concerning behavior; the composite scores based on previously validated scoring methods enable one to derive objective and verifiable results from the answers to the 125 questions.

The Data for the Twin Study

Responses to the 125 questions in the Bernreuter Personality Inventory were secured from all twins who could be located through a careful survey in the California high schools of San Jose, Palo Alto, Redwood City, and San Francisco. The group included 40 identical pairs, 44 like-sexed fraternal pairs, and 34 unlike-sexed pairs, or a total of 236 subjects. The ages of these children ranged from 12 to 19 years, with a mean of 16.2 and a standard deviation of 1.6 years. In addition, fifteen pairs of mature identical (monozygotic) twins were tested; these ranged in age from 20 to 65 years. The classification of twins into identical and fraternal (dizygotic) groups was made according to the techniques described by Bonnevie,[8] Dahlberg,[9] Siemens,[10] Newman,[11] Muller,[12] and others. In certain other reports, the writer[13] has indicated in greater detail the data considered in such classifications.

The Bernreuter Inventory is easy to administer, and usually does not require more than a half-hour of the subject's time. Since it is not intrinsically difficult, and since it gives subjects an opportunity to report about themselves, the inventory usually meets with little re-

[7] R. G. Bernreuter, *Manual for the Personality Inventory.*

[8] K. Bonnevie, "Studies of Papillary Patterns of Human Fingers," *Journal of Genetics,* 1924, XV, 1-111.

[9] G. Dahlberg, *Twin Births and Twins from a Hereditary Point of View* (Stockholm, 1926), pp. 85 and 296.

[10] H. W. Siemens, "The Diagnosis of Identity in Twins," *Journal of Heredity,* 1927, XVIII, 201- 209.

[11] H. H. Newman, "Studies of Human Twins: I. Methods of Diagnosing Monozygotic and Dizygotic Twins," *Biological Bulletin,* 1928, LV, 287-297.

[12] H. J. Muller, "Mental Traits and Heredity," *Journal of Heredity,* 1925, XVI, 433-448.

[13] H. D. Carter, "Identical Twins Reared Together," *Journal of Heredity,* 1932, XXIII, 53-66; and "Case Studies of Mature Identical Twins," *Journal of Genetic Psychology,* 1934, XLIV, 154-174.

sistance from unselected persons. In this study, testing was done either individually or in small groups. Preliminary explanations of the purpose of the general twin study, and a report of results, were made for the purpose of increasing coöperation. The younger subjects in the groups were given additional supervision and explanation where necessary.

The data for this report are the responses of the twins to the individual questions of the inventory, and the scores of the twins on each of the specialized scoring schemes of the inventory.

Effects of Age, Sex, and Intelligence

Scores upon this test are practically independent of age,[14] for the range of ages included in this sampling. The only scales which show any appreciable correlation with age are those for measuring dominance (r is .15) and sociability (r is .24). There is apparently a very slight tendency for the older children to be more dominant and *less* sociable.

Likewise, there is very little correlation between intelligence and scores on this inventory. Intelligence measures were available for only a portion of the sampling, consisting of 67 girls and 81 boys; the I.Q.'s ranged from 73 to 140, with a mean of 102.7 and a standard deviation of 13.8 points; this sub-group is probably fairly representative of the total group. The highest correlations with intelligence are those for self-sufficiency (r is .10) and sociability (r is .13). The available evidence therefore suggests that scores on this test are almost completely independent of intelligence.

Computations based upon the twin data of Table I showed large and reliable sex differences in raw scores. Girls are more neurotic, less self-sufficient, more introverted, less dominant, less self-confident, and more sociable than boys. All the differences save the last mentioned are statistically reliable. Thus certain well-known tendencies toward sex differences are confirmed when the scores of this group of twins are examined.

In the further treatment of the data, care has been taken to achieve results not dependent upon differences in age or sex. When percentile scores have been used, they have been based upon separate

[14] Some of the results in this section are taken from an earlier publication (H. D. Carter, "Twin Similarities in Personality Traits," *Journal of Genetic Psychology*, 1934, XLIV, 154-174), in which effects of age and intelligence were discussed. That earlier publication dealt with only a portion of the data. Twin-similarities were considered, but only four of the scales were available. Two of the questions raised in the present paper were not considered in the earlier report.

norms for the two sexes; likewise, the standard scores employed have been determined separately for males and females, to eliminate sex differences. In several instances the data from groups of males and females have been analyzed separately. Care has been taken to give independent treatment to results from the group of high-school students, in order that age effects or other factors associated with the selection of the other subjects might not affect the findings.

With sex differences taken into account, scores falling in all parts of the standard score range are found in each of the several sub-groups of twins. Apparently the tendency to secure high, average, or low standard scores is an individual matter, very largely independent of age, sex, or intelligence. Even if some relationships of scores to these variables may be assumed a priori, treatment of the data has been so arranged that conclusions are evidently not dependent upon factors associated with age, sex, or intelligence of the subjects.

Twins Compared with Non-Twin Groups

Since the present study is based upon data from a group of 266 persons who are twins, it is instructive to compare scores with those to be expected from a similar non-twin population. First, comparisons are made by referring the scores of the twins to Bernreuter's percentile norms. Accepting the test as valid, it appears that the twins are emotionally more stable than people in general; they are at the same time less self-sufficient, more extroverted, and about average in dominance. This general picture is characteristic of each of the sub-groups, with two exceptions; first, the identical (monozygotic) male twins and the fraternal (dizygotic) female twins seem to be more dominant than the other groups; second, the older identical twins seem to be somewhat different, being more unstable and introverted. Since the sub-groups are rather small, these specific sub-group differences may not be important, and should not be overemphasized.

More satisfactory comparisons can be made on the basis of means and standard deviations.[15] The results are given in Tables I and II, for the total groups of high-school age, for males and females separately. These results indicate that twins as a group are more than ordinarily stable and extroverted; the differences are statistically reliable for males taken alone, and the results for females are in the same direction. Norms for the Flanagan scales are avail-

[15] R. G. Bernreuter, *Manual for the Personality Inventory.*

able for males only; statistically reliable differences there indicate that male twins are more self-confident and more sociable than the norm populations.

Inter-comparisons of all the sub-groups of twins are not shown in Tables I and II, because the groups are not sufficiently large to establish important differences. Such comparisons were carried out, using the standard scores based on the total group of twins. It was found that the mature identical females are more neurotic than the high-school females, but the difference between means was only 1.6 times its standard error; the same ratio was found in comparing these two groups on introversion. The mature female group is also less dominant than the high-school group, the difference being 1.4 times its standard error. These differences, like all the differences between sub-groups, are only suggestive because of the small size of the populations compared.

Intercorrelations of the Traits

A better understanding of the nature of the several traits can be had through study of their intercorrelations. Because of the selective factors implied above, it has been considered worth while to compute the intercorrelations between the traits, for this group of twins, for comparison with similar results from other populations, as published in Bernreuter's manual.[16] The correlations, computed separately for males and females, are presented in Table III. Two of the intercorrelations are different from those for the norm population as found in the manual. In the twin material, submissiveness (lack of dominance) is less related to neurotic tendencies and introversion than in Bernreuter's population (college group). The age range of the present group is not sufficient for thorough analysis to prove that the youth of these subjects has not caused the difference; however, Flanagan's table[17] of correlations based upon data from high-school boys permits direct comparison. Flanagan's correlations based upon 305 eleventh-grade boys are –.80 for neurotic tendencies with dominance, and –.69 for introversion with dominance; these correlations are higher than for our twins. For some reason, in the present group of subjects, submissiveness is apparently less related to neurotic tendencies than in other comparable populations. In explanation, it might be suggested that twins often learn acceptable,

[16] *Ibid.*

[17] J. C. Flanagan, *Factor Analysis in the Study of Personality* (Stanford University, Calif., Stanford University Press, 1935).

coöperative attitudes as a solution of face-to-face rivalry situations. Among ordinary siblings, one dominates the other more often, causing the presence, in a large group, of many individuals with submissive attitudes accompanied by distress and maladjustment. Among twins, if one is submissive the other is also likely to be (see Table VII); such submissiveness may occur along with good adjustment to social situations.

Twin-Similarity in Answering Inventory Questions

It can be ascertained through a general type of analysis whether identical (monozygotic) twins tend to answer these inventory questions more similarly than do fraternal (dizygotic) twins. Casting aside the particular scoring schemes, one might be inclined to ask whether there is a tendency for both members of twin pairs to say *Yes,* or for both to say *No,* more often than chance would explain. That is, if one twin says *No* in response to a question (any question), is there a tendency for the other twin to say *No* more often than chance, in response to that same question?

Two types of analysis will be described. In the first, the nature of the particular questions is disregarded; we are considering for the moment merely the activity of answering the rather personal questions of the inventory. In studying this form of behavior, we wish to know whether twins give similar answers more often than one could expect according to chance, and whether the similarity indicated is greater for identical than for fraternal twins. All questions are lumped together for consideration of general findings. Here we should note Bernreuter's statement that "An inspection of the individual items on the tests shows that most of them refer principally to emotional aspects of behavior." In lumping together the results of 125 different questions, we are considering for the moment twin-similarities in emotional behavior as revealed by standardized, independently made self-reports.

The responses given by each pair of twins to each of the 125 questions were tabulated in such a way as to indicate the number of times both twins answered *Yes,* or *No,* or question mark (three kinds of identical answers), the number of times one twin responded with yes, while the other replied no (opposite answers), and the number of times one replied *Yes,* or *No,* and other replied question mark (two ways of indicating intermediate amount of disagreement). When results for many pairs of twins were incorporated in

one tabulation, group comparisons based upon thousands of responses were made possible. Such tabulations were analyzed to determine for each group the extent to which agreement was greater than chance; a measure of resemblance based upon elementary probability (the coefficient of mean square contingency) was thus secured.[18] The possibility of regarding *Yes*, *?*, and *No* as ordered categories also merits consideration; therefore measures of product-moment correlation were computed for these same data by means of an appropriate formula devised by Pearson.[19]

The quantitative results of this analysis are presented in Table IV. Since the responses to 125 questions are considered collectively, thousands of responses are involved in each of the tabulations. This results in rather stable coefficients of resemblance, which must of course be interpreted in relation to the procedure. The similarity of responses is greatest for identical twins (r is .51), intermediate for like-sexed fraternal twins (r is .37), and least for unlike-sexed twins (r is .26). Female pairs respond more similarly than do males. In the case of females, but not males, older pairs respond more similarly than do younger pairs. These findings reflect intra-pair similarities, in responses to questions concerning customary behavior in emotionally charged situations. This treatment of results has been supplemented by consideration of similarity in responses to particular items.

Item-Analysis

Tabulations were made for each individual question, showing the frequency of similar and dissimilar responses by the twins. This was done for each group separately, so that results from identical and fraternal groups could be compared. This treatment of the data should show whether the greater general similarity of identical twins is apparently due to similarity in answering of particular types of question. It should also show whether there are any types of question which are answered more similarly by the fraternal twins.

Table V presents a selection of questions answered most similarly by identical twins as compared with like-sexed fraternal twins. Table VI shows the selection of questions which were answered most similarly by fraternal twins of like-sex as compared with monozygotic twins. In selecting these questions from the total 125, a simple

[18] See K. Holzinger, *Statistical Methods for Students in Education* (Boston, Ginn and Co., 1928), for the statistical procedures involved.

[19] K. Pearson, "On the Measurement of Broad Categories upon Correlation," *Biometrika*, 1913, IX, p. 116.

measure was used instead of the correlation method applied to group results; this simplification seems preferable in view of the nature of the material and the labor involved in the alternative procedure. The percentage of similar answers, minus the percentage of opposite answers, was computed for each group of twins, for each question. The resulting figure for fraternal twins was subtracted from that for identical twins, to yield the final measure for each question. This final figure, if positive, indicates greater similarity of the identical twins.

These final measures varied considerably for different questions; hence a selection was made to reduce the bulk of the data. The mean and standard deviation were computed for the 125 measures pertaining to the 125 questions; the mean was positive, indicating greater general similarity of identical twins. The final measures were expressed in standard score form, and those more than one S.D. above and below the mean were selected for presentation in Tables V and VI. For each question, the standard score indicating its position in the series of questions is given in the table. Table V therefore lists questions with high positive values (identical twins more similar), and Table VI lists questions with extreme negative values (fraternal twins more similar). The sigma values of the questions are given in parentheses for the reader's convenience, although the writer feels that results secured here from single questions are sufficiently unreliable to make these standard scores of distinctly limited value.

Inspection of Tables V and VI leads to no clear-cut explanatory generalizations. If there is any distinctive quality characteristic of all the questions of either group, it is not apparent to the writer. Probably a repetition of this procedure with another similar group of twins would not show agreement of details of the sort given in the item-analysis. This item-analysis was carried out for males and for females separately, but in view of the inconclusive nature of such results, it does not seem worth while to present the additional tabular material. The unreliability of item-results based on small groups renders such subdivision of material quite unproductive. It is possible, perhaps even likely, that similar item-analysis based on a *much larger* group would lead to significant discoveries.

Apparently there is a general similarity of identical twins in answering inventory questions; this general similarity is greater than that for fraternal twins. Measures of this similarity based on

single questions are unreliable; such measures fluctuate so much that any important underlying factors are likely to be obscured. The method of item-analysis seems a priori to be valuable, but present findings indicate that, if clear-cut results are to be found, a much larger body of data is required. At present, application of the several scoring schemes for the Bernreuter Test, to secure measures based upon all the questions, seems more productive as a method of discovering trends underlying twin-similarities.

Self-Confidence and Sociability

In considering twin-similarities, emphasis is here placed upon the two new scales (Flanagan) measuring self-confidence and sociability; this is done because the results are new, and because the scales have been developed by a method which results in independent measurements. In Table VII the correlations are presented, along with similar results for the other four traits, taken from a previously published table.[20] Because of the small size of the sub-groups, results are shown only for the combined groups of males and females.

It is apparent that identical twins are most similar, like-sexed fraternal twins are second, and unlike-sexed twins are least similar. These results are not uniform for all the traits; a noteworthy exception is the dissimilarity of fraternal like-sexed twins in self-sufficiency. Differences in the reliability of different scales can hardly account for the variation in coefficients, since the reliability coefficients are all very much alike;[21] errors due to small sampling could conceivably account for the differences, but such explanation is improbable. Identical twins are much more similar in dominance than in introversion, and the other correlations are intermediate. Correlations computed separately for males and females were very similar for identical twins, but fluctuated more in the case of fraternal twins, probably because of unreliability of low correlations based on small groups. The average coefficient is .58 for identical twins (high-school group only), .26 for like-sexed fraternal twins, and .19 for unlike-sexed twins.

Nature or Nurture?

Results such as these are consistent with hypotheses favoring explanation of twin similarities in terms of either nature or nurture.

[20] H. D. Carter, "Twin Similarities in Personality Traits," *Journal of Genetic Psychology*, 1933, XLIII, 312-321.

[21] R. G. Bernreuter, *Manual for the Personality Inventory*.

If the traits here examined are dependent primarily upon inherited mechanisms, one would expect results such as have been found; for example, if self-confidence is dependent largely upon inherited physical or mental characters, monozygotic twins should be more similar than fraternal twins in self-confidence. Likewise, if the traits are largely developed through environmental influences, one should expect results to be as indicated in this study. This is true because the environments of identical twins are more alike than those of fraternal twins; this fact is perfectly obvious to persons who have had intensive field contact with twin groups;[21a] furthermore, it is supported by Wilson's study,[22] in which objective aspects of environment were investigated in relation to twin resemblances.

Such studies as the present one do not untangle nature and nurture effects, but they contribute information about twin resemblance, and they may aid in indicating the emphasis to be placed upon each set of influences, for different traits. Such studies, combined with experimental investigations, may ultimately show which types of experience seem to affect development of particular attitudes. It has been shown that traits of the sort here investigated have little relationship to age, and that coefficients of twin resemblance are appreciably lower for attitudes than for abilities,[23] although it is commonly assumed that attitudes are more affected by training. Furthermore, case studies indicate that twin differences are not consistently related to the environmental differences studied.[24] From these facts, one is inclined to the opinion that underlying hereditary factors are more important in the development of attitudes than might at first be supposed. An alternative explanation is that the development of attitudes of the sort here studied is conditioned largely

[21a] Fraternal twins, having greater physical differences, often cannot do the same work, or play the same games. Likewise, being different in intellectual ability, they do not make the same progress in school. Finally, being different in interests, they do not choose the same amusements. The intra-pair differences are usually greater for fraternal twins, and they lead to greater differences in experience. Some of the differences, such as in physique, are often clearly of hereditary origin; hence we may infer that hereditary differences lead to environmental differences.

[22] P. T. Wilson, "A Study of Twins with Special Reference to Heredity as a Factor Determining Differences in Environment," *Human Biology*, 1934, VI, 324-354.

[23] H. D. Carter, "Twin Similarities in Occupational Interests," *Journal of Educational Psychology*, 1932, XXIII, 641-655; and "Twin Similarities in Personality Traits," *Journal of Genetic Psychology*, 1933, XLIII, 312-321.

[24] H. D. Carter, "Identical Twins Reared Together," *Journal of Heredity*, 1932, XXIII, 53-66; and "Case Studies of Mature Identical Twns," *Journal of Genetic Psychology*, 1934, XLIV, 154-174.

by environmental factors other than the obvious ones associated with economic status, formal education, etc. If environmental influences are operative, it seems likely that they are subtle influences associated with social life, the family, the personal problems of the individual, and the general *experienced* environment of the individual. The influences operating must be those which are more similar for identical than for fraternal twins—a statement which potentially has much meaning.

Summary and Conclusions

The Bernreuter Personality Inventory has been administered to 133 pairs of twins, and the results have been analyzed to study the twins with respect to emotional stability, self-sufficiency, introversion, dominance, self-confidence, and sociability, as measured by the six scoring scales which have been devised for this inventory. In addition, a more general analysis has been carried out, to investigate similarity in answering of the questions, without use of the specialized scoring schemes; this was supplemented by an item-analysis. The following conclusions are indicated by the data:

1. Twins appear to be more stable, more extroverted, and more self-confident than children in general. They are also more sociable and gregarious. There appears to be no reliable difference in dominance or in self-sufficiency.

2. Identical twins are more similar than fraternal twins, in all the traits considered, except possibly introversion. Like-sexed fraternal pairs tend to be more similar than unlike-sexed pairs, except in self-sufficiency and sociability.

3. Identical twins give similar answers to these personal inventory questions more often than do fraternal twins. Thus it is to be expected that almost any scoring scheme which makes use of a large number of the questions will indicate greater similarity of identical twins in the traits measured.

4. Item-analysis indicated that fraternal twin pairs *of these groups* did answer *some* questions more similarly than the identical twin pairs. A possible explanation may be made in terms of unreliability of individual item responses; that is, we may postulate a greater general similarity of identical twins, which is revealed adequately only by the results based upon a large number of questions.

5. These results are interpreted as evidence of general similarity of twins with respect to emotional aspects of customary behavior; the information is secured by means of standardized, independent self-reports.

TABLE I. MEANS AND STANDARD DEVIATIONS OF DISTRIBUTIONS OF SCORES FOR TWINS AND FOR HIGH-SCHOOL NORM GROUPS

	MALES			FEMALES		
	n	Mean	S.D.	n	Mean	S.D.
Neurotic Tendencies						
Twins	128	—89.3	63.5	108	—44.8	83.7
Norm Group	145	—60.2	75.4	121	—22.9	78.2
Self-Sufficiency						
Twins	128	15.9	33.9	108	— 5.4	42.4
Norm Group	186	19.5	48.4	160	— 2.4	56.6
Introversion						
Twins	128	—57.7	36.7	108	—27.3	53.3
Norm Group	187	—32.5	46.6	160	—11.4	47.6
Dominance						
Twins	128	40.0	58.4	108	10.7	64.3
Norm Group	186	38.5	54.8	160	10.3	59.6
Self-Confidence						
Twins	128	—42.0	70.8	108	10.2	91.6
Norm Group	200	—14.5	92.8	...	....	....
Sociability						
Twins	128	—50.3	58.9	108	—66.9	74.6
Norm Group	200	—16.1	62.9	...	....	....

The statistics for the norm group are taken from the new norm sheet, published by Stanford University Press, 1935. In connection with the last two scales, note that positive scores indicate lack of self-confidence and lack of sociability respectively.

TABLE II. INDICATION OF THE STATISTICAL RELIABILITY OF THE DIFFERENCE IN SCORES BETWEEN TWINS AND NON-TWINS

	Difference	S.E. Difference	Difference / S.E. Difference
Neurotic Tendencies			
Males	29.07	8.41	3.46
Females	21.89	10.74	2.04
Self-Sufficiency			
Males	— 3.61	4.64	.78
Females	— 2.95	6.05	.49
Introversion			
Males	25.16	4.71	5.35
Females	15.87	6.36	2.50
Dominance			
Males	1.50	6.54	.23
Females	.38	7.78	.05
Self-Confidence			
Males	—27.53	8.70	3.16
Sociability			
Males	—34.24	6.85	5.00

The positive differences indicate higher scores for the twins than for the non-twin populations.

TABLE III. INTERCORRELATIONS OF BERNREUTER TEST SCORES. VALUES ABOVE THE DIAGONAL ARE THOSE COMPUTED FOR 128 HIGH-SCHOOL BOYS (TWINS); VALUES BELOW THE DIAGONAL ARE THE CORRELATIONS FOR 108 HIGH-SCHOOL GIRLS (TWINS)

	B1-N	B2-S	B3-I	B4-D	F1-C	F2-S
B1–N	...	—.41	.91	—.66	.85	.32
B2–S	—.34	...	—.32	.57	—.56	.42
B3–I	.96	—.23	...	—.56	.80	.36
B4–D	—.76	.51	—.69	...	—.84	.28
F1–C	.92	—.48	.88	—.84	...	—.02
F2–S	.46	.45	.51	.08	.22	...

The explanation of the symbols is as follows: B1-N, Neurotic Tendency; B2-S, Self-Sufficiency; B3-I, Introversion; B4-D, Dominance; F1-C, Self-Confidence; F2-S, Sociability.

TABLE IV. COEFFICIENTS OF MEAN SQUARE CONTINGENCY, AND COEFFICIENTS OF CORRELATION BY PEARSON'S METHOD FOR BROAD CATEGORIES, INDICATING THE EXTENT TO WHICH TWINS ANSWER INVENTORY QUESTIONS SIMILARLY

	Number of Responses	Raw C	Corrected C	r
Unlike-Sexed Twins				
34 Pairs	8500	.22	.31	.26
Fraternal Twins				
22 Pairs Males	5500	.28	.38	.33
22 Pairs Females	5500	.34	.47	.42
Male Identical Twins				
5 Older Pairs	1250	.31	.45	.47
25 Younger Pairs	6250	.38	.52	.50
All 30 Pairs	7500	.36	.51	.49
Female Identical Twins				
10 Older Pairs	2500	.42	.63	.66
15 Younger Pairs	3750	.39	.56	.53
All 25 Pairs	6250	.41	.59	.58
Averages:				
All Fraternal Like-Sexed Twins	....	.31	.42	.37
Younger Identical Twins	....	.38	.54	.51
Older Identical Twins	....	.37	.54	.56

TABLE V. QUESTIONS WHICH IDENTICAL TWINS ANSWERED MUCH MORE SIMILARLY THAN DID FRATERNAL LIKED-SEXED TWINS*

Are you often in a state of excitement? (3.0)
Does your mind often wander so badly that you lose track of what you are doing? (2.2)
Do you usually feel a great deal of hesitancy over borrowing an article from an acquaintance? (2.1)
Are you touchy on various subjects? (1.9)
Have books been more entertaining to you than companions? (1.8)
Would you "have it out" with a person who spread untrue rumors about you? (1.7)
Do you prefer to be alone at times of emotional stress? (1.6)

*The figures in parentheses indicate the critical ratios of the differences between identical and fraternal twins.

TABLE V (Continued)

Do you often experience periods of loneliness?	(1.5)
Are you greatly embarrassed if you have greeted a stranger whom you have mistaken for an acquaintance?	(1.4)
Do you usually object when a person steps in front of you in a line of people?	(1.4)
Do you have difficulty in starting a conversation with a stranger?	(1.4)
Do you feel that marriage is essential to your present or future happiness?	(1.3)
Do you like to be with people a great deal?	(1.2)
Do you prefer making hurried decisions alone?	(1.2)
If you are spending an evening in the company of other people, do you usually let someone else decide upon the entertainment?	(1.2)
Do you find it difficult to speak in public?	(1.2)
Do you often feel just miserable?	(1.1)
Do you try to get your own way even if you have to fight for it?	(1.1)
Are you inclined to study the motives of other people carefully?	(1.1)
Are you very talkative at social gatherings?	(1.1)
Are you troubled with shyness?	(1.0)
Do you take the responsibility for introducing people at a party?	(1.0)

*The figures in parentheses indicate the critical ratios of the differences between identical and fraternal twins.

TABLE VI. QUESTIONS WHICH FRATERNAL LIKE-SEXED TWINS ANSWERED MORE SIMILARLY THAN DID IDENTICAL TWINS

Do you consider yourself a rather nervous person?	(—3.0)
Do you worry too long over humiliating experiences?	(—2.2)
Do you ever argue a point with an older person whom you respect?	(—1.9)
Are your feelings easily hurt?	(—1.8)
Does admiration gratify you more than achievement?	(—1.6)
Do you daydream frequently?	(—1.5)
Are you careful not to say things to hurt other people's feelings?	(—1.5)
Do you frequently argue over prices with tradesmen or junkmen?	(—1.4)
Do you blush very often?	(—1.4)
Have you ever organized any clubs, teams, or other groups on your own initiative?	(—1.4)
Do you tend to be radical in your political, religious, or social beliefs?	(—1.3)
Are you able to play your best in a game or contest against an opponent who is greatly superior to you?	(—1.2)
Do you usually avoid asking advice?	(—1.2)
If you were hiking with a group of people, where none of you knew the way, would you probably let someone else take the full responsibility for guiding the party?	(—1.2)
Are you troubled with the idea that people on the street are watching you?	(—1.2)
Does discipline make you discontented?	(—1.1)
Are you much affected by the praise or blame of many people?	(—1.1)

TABLE VII. TWIN-SIMILARITIES IN SELF-CONFIDENCE (F1-C), SOCIABILITY (F2-S), NEUROTIC TENDENCIES (B1-N), SELF-SUFFICIENCY (B2-S), INTROVERSION (B3-I), AND DOMINANCE (B4-D)

	B1-N	B2-S	B3-I	B4-D	F1-C	F2-S
Identical Twins:						
40 Young Pairs	.61	.59	.43	.75	.56	.54
All 55 Pairs	.63	.44	.50	.71	.58	.57
Fraternal Twins:						
44 Like-Sexed Pairs	.32	—.14	.40	.34	.20	.41
34 Unlike-Sexed Pairs	.18	.12	.18	.18	.07	.39

NEWS AND NOTES

Critical Comment

The Use of the Probable Error in Evaluating Clairvoyance

A recent article by J. B. Rhine contains the following reference to his mathematical procedure, which we find clearer than anything on the topic previously published by him: ". . . significantly high, that is, whose total of successes surpassed the mean chance expectation by a deviation more than four times the probable error of the mean. . . . The . . . probable error . . . is found by the formula p.e. $= .6745\sqrt{npq}$, where n is the number of trials, p the probability of success per trial, and q the probability of failure. When the p.e. is found, it is simple to state how many times the p.e. the deviation is, and this value (X) is convertible by the use of normal probability integral tables (when n is large, as in all our work, the curve approximates that of a normal distribution and skewing may be neglected) into P or the probability of the results occurring by chance alone."[1]

Rhine's situation may be represented by a coördinate system of which the abscissa indicates number of successes (increasing from left to right) and the ordinate indicates frequency. To the left may appear a vertical line marking the location of the theoretical chance mean np; to the right, a similar vertical line marking the location of the number of successes scored by the subject. These are the observed data. Back of the left-hand or np line, however, we must imagine the theoretical chance distribution of which the np line is the mean; and since, in the case under discussion, p is 1/5 and q is 4/5, this distribution is necessarily skewed, with the longer tail toward the right. Back of the right-hand line (the subject's score), likewise, we must imagine a corresponding distribution of a large number of guesses by the same subject under the same conditions.

In any case, the variability of this last distribution will not be satisfactorily given by the formula quoted, since there will be characteristic individual factors not taken account of by its assumptions; but more important is the fact that, since p and q have the same asymmetrical values as before, this distribution will—if nothing but chance operates—also be skewed. Thus if we assume (the most probable case) that the observed performance of the subject lies at the mean of his infinite distribution, we have two distributions, both skewed upward, and the problem is that of determining the proportion of cases in which it is to be expected that the means will lie as far apart as they are observed to lie in the case observed.

Rhine has evidently proceeded as if the appropriate measure of this distance were the variability of the *subject's* distribution, which he has (as we believe erroneously) equated with the formula. The implication would seem to be that the datum is the subject's performance, from which the theoretical mean is a deviation whose significance is to be measured (unless a different concept be indicated by the last three words of the first clause quoted). The real state of affairs would seem to be the

[1] This journal, 1934, III, p. 97.

reverse; i.e., the theoretical mean is the norm from which the deviation of the subject is to be measured. In any case, it seems clear that normal probability tables are not applicable to the evaluation of differences between values pertaining to skewed curves, and that, since the curves are theoretically and necessarily skewed, the multiplication of cases will not remedy the situation. The correct evaluation of the case seems to call for comparison of areas under skewed frequency-curves of Pearson's Type IIIi.

RAYMOND ROYCE WILLOUGHBY.

A Reply to Dr. Willoughby

R. R. Willoughby criticizes the mathematical procedure used by J. B. Rhine on three main points, which, taken in reverse order, are: (1) that "Rhine has evidently proceeded as if the appropriate measure of this distance [the deviation from np] were the variability of the *subject's* distribution. . . ." As the only measure of variability used in the Rhine reports to date has been P.E. $= .6745\sqrt{npq}$ which is obviously derived from the variance of a theoretical binomial series, the grounds for the foregoing charge are far from "evident" as is stated.

(2) The criticism is offered that subject variation will not be given satisfactorily by the Rhine formula "since there will be characteristic individual factors not taken account of by its assumptions." This clearly grants the conclusion sought. "Characteristic individual factors" are not "chance" factors. If they appear, then it is necessary to conclude that a chance hypothesis is insufficient to account for the subject's performance.

(3) The foregoing depends, however, in part upon the main criticism offered: that the Rhine use of the normal curve for evaluation is faulty because "the theoretical chance distribution of which . . . p is 1/5 and q is 4/5 . . . is necessarily skewed" such that "correct evaluation seems to call for comparison of areas under curves of Pearson's type IIIi." This is a clear matter of fact and may be demonstrated conclusively one way or the other.

The chance distributions referred to must be of the order $(1/5 + 4/5)^n$. If n is taken as 25 (the smallest feasible number of guesses in the Rhine experiments, though no report was ever made of so small a number) the skewness is given by

$$sk = \frac{\sqrt{\beta_1}\,(\beta_2 + 3)}{2\,(5\beta_2 - 6\beta_1 - 9)} \pm \chi_1 \sqrt{n}\ \Sigma sk;$$
$$= .1636 \pm .1619$$

$$\beta_1 = .09 \pm .189$$
$$\beta_2 = 3.01 \pm .777$$

which indicates that for $n = 25$ the skewness of the distribution is not significantly different from zero. Since the skewness and β statistics are all insignificantly different from the Gaussian $sk = 0, \beta_1 = 0, \beta_2 = 3$, the distribution may be regarded as *normal*. The skewness and β statistics of the distributions under discussion all approach the Gaussian as n increases; so the larger n becomes, the better is the fit to a normal curve. With the normal curve adequate to describe the expected distributions there is no need to fit a type III curve to them.

It must be concluded that the critical comments offered by Willoughby are entirely inapplicable to the Rhine procedure.

C. E. STUART.

U. S. A.

A Sign of the Times

As psychology has been treated somewhat offishly by the physical sciences since its inception, we may view with interest the development of

the controversy between Einstein and Niels Bohr on the impossibility of describing physical reality in the light of the quantum theory, because in the process of being applied, the criterion is no longer the same absolute standard that it was *in abstracto*. We may remember that the great stumbling-block of psychology in its introspective phase had been, as James so picturesquely described it, quickly turning out the gas so as to see what it looks like to be in the dark or, as someone else expressed it, "trying to catch the shadow." Physics, we note, is not immune from the so-called psychological limitation.

"Effective Temperature"

In a similar connection we find that a good part of what is called the heat wave, and what is described generally as humidity, belongs to the mental sphere. Of course we have always known that our attitude is often responsible for the "feeling" of heat or cold, but when a meteorologist, Dr. D. R. Morris, of the New York Meteorological Branch of the Weather Bureau in Central Park, declares that "effective temperature includes the psychological effect which involves the all-important time factor of the duration of the heat wave and the longer the wave lasts the worse the psychological effects," we begin to find official confirmation of what psychologists have been suspecting for decades.

A. A. A. S. Meeting

At the meeting of the American Association for the Advancement of Science, which was held in Minneapolis during the last week in June, a number of papers were devoted to psychology. Perhaps it was due to the environment that for the most part the papers dealt with the psychology of learning and child psychology. J. E. Anderson, speaking on "limits of learning," reviewed the literature on maturation and learning done in young children, and on this basis suggested the importance of analysis in terms of complexity of the task set.

H. A. Sorenson showed that adult learning preserved the ability acquired in early days more considerably than had been supposed.

At a joint session of the psychological and educational sections Jean Marques discussed the development of logical processes in school children and concluded that there was an increase in causal reasoning with age and school grade, although not with socio-economic status. Piaget's classification in this connection was not found to hold, and another classification was proposed.

In the symposium on "attitude measurement," E. A. Rundquist brought out the significance of the form of statement in personality-scale construction. Statements made in negative form give different results from those made in positive form, a conclusion by the way which has already been reached by investigators (e.g., Langfeld, *et al.*) in other spheres. R. F. Sletto argued for the inadequacy of internal consistency as a criterion in scale construction.

Psychology in Sunday Newspapers

The Ishihara test for color blindness, which is sold by the C. H. Stoelting Co., has apparently come to the attention of the magazine editor of a chain newspaper and has been reproduced in color for the benefit of the readers. This brings to mind the necessity of copyrighting such material in other countries besides the one where it may be originally published.

An International Psychodiagnostic Society

There is a movement on foot to organize a society for the study of psy-

chodiagnostics (embracing such behavior as gait, voice, handwriting, gestures, etc.) along strictly scientific lines. Only those who have already made contributions to the subject will be invited to join. While the society will be established primarily in the United States, it will have an international membership and status. All psychologists interested in this field of personality may communicate with A. A. Roback, 8 Prescott St., Cambridge, Massachusetts, for more information.

Education in the United States

While the military program of the country has been expanded in the interest of preparedness for war, despite the economic condition of the country, it has been established that one-eighth of the school children in the United States are in school districts without sufficient funds to operate the customary term.

Psychological Department at Harvard University

The Psychological Department at Harvard University will undergo certain changes, beginning next fall. Professor Karl S. Lashley, of the University of Chicago, who has made notable contributions to the field of animal psychology and particularly to that of cerebral functions, has been appointed professor of psychology. His laboratory will be situated in the biological building. Psychology at Harvard University will thus be scattered over the following locations: Emerson Hall (main experimental laboratory), Psychological Clinic (abnormal psychology), Psycho-educational Clinic (educational psychology). Boylston Hall (animal psychology), School of Business Administration (business and industrial psychology), and the Biological Institute (cerebral functions).

A. A. Roback.

Great Britain

The Heath Clark Lectures on "Temperament"

In view of the marked present-day interest in the investigation of the conative side of human nature, following the fruitful work that has been done on cognition, the subject arranged by the National Institute of Industrial Psychology for the 1935 Heath Clark Lectures—"The Temperamental Factor in Industry"—was a particularly happy one.

On May 13 and 20, Professor Major Greenwood spoke on "The Modern Importance of the Temperamental Factor and Its Ancient History"; and on "Temperaments Physical and Psychological in Modern Science"; and on May 27, Dr. May Smith completed the course by dealing with "Industrial Applications."

In his first lecture, Professor Greenwood showed how some ancient teaching on temperament has its modern application, while some modern speculation is but ancient error in a new dress. The practical side of Galen's psychology of the second century, stressing as it did the fundamental distinction between the cognitive and affective processes, is of living interest. Though his doctrine of temperaments seems at first absurd, it is less dangerous than that of some modern exponents of endocrine dogmas, inasmuch as Galen's classification, however trivial, was recognized by him as just a classification, while the new endocrine labels suggest that we know something we do not know.

Where the ancient teaching on temperament is unsatisfying is in its vagueness on the side of prognosis. How far modern research has corrected such vagueness furnished the basic question for Professor Greenwood's second lecture. A review was given of the work of various modern investigators, including the anatom-

ical descriptions of the Italian school and of Kretschmer, the rating methods of American psychologists, and the mathematical procedure of the Spearman school. Special emphasis was laid on the importance of the fact that, fundamentally, statistical description is only *group* description, and the conclusion was reached that it may soon be possible to make a rational prognosis of the fate of *groups,* based on temperamental assessment, but as regards individual prognosis, though we may perhaps be on the eve of great advances, we are not at present able to replace the artist by the scientist.

The third lecture on "Industrial Applications," delivered by Dr. May Smith, provided an appropriate underlining of Professor Greenwood's final point concerning psychology as an art, for it was in effect (although not so described) an account of actual approaches by "scientific artists" to the problems of temperament in industry. The account was enriched by incidental historical illustrations and by an illuminating discussion of the qualities desirable in "captains of industry" and their lieutenants; the lecture concluded with a plea for more detailed study of the problem of "industrial misfits."

G. W. GOODALL.

The Effect of Localized Brain Injuries on Handwriting

In the *Schweizer Archiv für Neurologie und Psychiatrie* (1934, XXXIV, 38-68), E. Feuchtwanger (Munich) and K. Román-Goldzieher (Budapest) describe 11 cases of localized cerebral lesion and try to ascertain whether and how the writing movement has been affected by such injuries. The authors claim not only that post-traumatic handwriting shows marked alterations, but that the alterations in the handwritings of patients with injuries to the frontal lobes tend to be different from those caused by lesions in the central and posterior parts of the brain.

Unfortunately in most of the cases discussed the time interval between the pre- and post-traumatic handwriting is very long (fifteen years or more), and further, several of the specimens previous to the injury were written at the time of puberty. No wonder that the difference between the pre- and post-traumatic handwritings are often so striking and that in cases of injuries to the frontal lobes they are attributed by the authors to changes in the emotional make-up! The authors seem, however, to hold that in normal individuals developmental changes in handwriting are much less marked—a claim which is not always borne out by facts.

As a case study the paper is an interesting contribution to the research on pathological handwriting. The authors are, of course, aware of the fact that the small number of cases does not allow of definite conclusions. They also endeavor to off-set the subjectiveness of their graphological observations by measuring writing pressure with a special instrument, the so-called graphodyn, invented by Dr. Román-Goldzieher.

JAN MELOUN.

Scandinavia

The *Sixth Scandinavian Congress of Psychiatry* took place in Stockholm, August 22 to 24, 1935, its main theme being the somatology of schizophrenia. A report on the proceedings will appear later in these columns.

The Swedish Medical Association for Psychotherapy

In the first three numbers of the Swedish medical journal *Swenska Läkareselskabets Förhandlingar* of this year, we find an interesting discussion as to the advisability of

forming, within the Swedish Medical Association, a special section for psychotherapeutic research.

A motion to this effect, which in Denmark was carried without difficulty, caused a heated discussion in the Swedish Association, because its more conservative members held that the interests of psychotherapy could be attended to by the already existing sections of psychiatry and neurology. In spite of the clever advocacy by Dr. Poul Bjerre of Stockholm, the well-known Swedish psychotherapist and president of the Swedish section of the international *Allgemeine aerztliche Gesellschaft fuer Psychotherapie*, the motion was defeated by sixty-one against twenty-five votes.

It is remarkable that the position of psychotherapy among the medical disciplines is still somewhat problematic, and that, after twenty-five years of unceasing effort, Dr. Bjerre, whose brilliancy is admitted by his opponents, did not succeed in gaining sufficient support for the formation of an autonomous section of psychotherapy. One of the reasons is undoubtedly the conniving and, in a few cases even friendly, attitude of many psychotherapists towards the layman-analyst, an attitude rightly regarded by most doctors as dangerous to the patients and incompatible with the dignity of the medical profession.

OLUF BRÜEL.

BOOK REVIEWS

THE RIDDLE OF THE SPHINX. By Geza Roheim. With a preface by Ernest Jones. London: The Hogarth Press and the Institute of Psycho-Analysis, 1934. Pp. 302. 18*s*.

It is possible to review this book in two ways. The reviewer may take the book as it stands, analyze the argument, examine the methodology by which the author arrives at his conclusions, test the validity of these conclusions and pronounce judgment upon *The Riddle of the Sphinx* in terms of the internal evidence alone. Alternatively, the reviewer may consider this book as one section of Dr. Roheim's considerable contribution in the field of the inter-relationships between ethnological and psycho-analytical material,[1] and it is not even necessary to stop there, but the argument of the book may be placed in the still wider context of the whole psycho-analytic ethnological discussion.

If the first method be followed, the verdict is inevitably negative. Dr. Roheim presents an argument bewildering in its tangle of unresolved complexities, elisions, and condensations. After a short introduction posing the problem of the origin of culture, he presents a long discussion of Australian aboriginal ceremonialism which is practically unintelligible in the terms in which he describes it here. If his previous publications on Australia have been read with great care, and if the reader possesses a good working knowledge of Freudian theory, and a detailed knowledge of the ethnology of Central Australian tribes drawn from the other more formal sources,[2] this section on Australia becomes relatively intelligible, and also very stimulating. In the present volume, however, it is interlarded with references to European folk-lore, on the basis of a frame of reference which is seldom more than implied. Dr. Roheim traces the latent content of Australian ceremonialism to what he considers to be the central trauma of the upbringing of a child among these people, the fact that the mother sleeps upon her growing son, so: ". . . the Aranda are a people who have repressed the '*alknarintja* situations' (this sleeping experience) and, for this reason, make a vagina on their penises (subincision) and decorate their *tjurungas* (penis symbol) with concentric circles, (vaginal symbols)" (p. 169). He presents dream material, linguistic material, ceremonial detail to support his view. He then presents a short chapter, far sketchier and even less clear than the Australian section, upon the Dobuan people of Normanby Island, where again it is necessary to invoke his previous publications on this culture and also Dr. Fortune's work,[3] in order to follow him at all. He sketches in a series of related myths, discusses the content of the form found in matrilineal Dobu where the central trauma is the

[1] "Psycho-Analysis of Primitive Culture Types," *International Journal of Psycho-analysis*, Australian Research No., Jan.-April, 1932; "Women and Their Life in Central Australia," *Journal of the Royal Anthropological Institute*, LXIII, p. 277.

[2] Sir Baldwin Spencer and F. J. Gillan, *The Arunta . . .* (2 vols.; London: MacMillan and Co. Ltd., 1927).
C. Strehlow, *Die Aranda und Loritja-Stamme in Zentral Australien*, I, "Die totemistischen Kulte der Aranda und Loritja-Stamme," 1910.

[3] R. F. Fortune, *Sorcerers of Dobu* (London: George Routledge and Sons, Ltd., 1931; New York: E. P. Dutton and Co., 1932).

mimicry of being eaten by the father in infancy followed by a suppression of the patrilineal principle in later life. With these myths are paralleled a series of myths of similar plot but which are selected from widely different culture about which Dr. Roheim has no supporting material whatsoever. In this discussion he unhesitatingly, on the basis of his Dobuan detailed analysis, identifies the cannibalistic monster as "parent of the growing hero," although the abstracts which he gives do not support this view, and ignores all of the careful work which has been done on the question of diffusion and incorporation of myths, in his statement "we can not explain how a figure in a foreign myth can have become a national hero." Therefore he assumes an original relationship between these myths in the remote past. This chapter is definitely a step backwards in the psycho-analytic use of myth, just as the Australian chapters, although unclear and confused, are a very definite step forward, because instead of regarding the myths as related to the living reality of the experience of recent generations, he uses them merely as clues to a hypothetical past.

Dr. Roheim next attacks the problem of human origins for which his analysis of present-day contemporary primitive culture has been merely a preparation, although he nowhere comes to grip with the problem of the validity of using contemporary primitive societies as accurate indicators either of the social structure of our own past, or indeed of those peoples' own past. He discusses Freud's theory of the primal horde and invokes primate material[4] which fits in best with theories of a dominant old male. (It would be well to place beside this discussion Ray Carpenter's study of the Howling Monkeys.[5]) He rephrases Freud's hypothesis so as to emphasize that it is a statement of a long-time span, in which out of the conflict between fathers and sons, culture was gradually evolved, and suggests that it is not merely the interjection of the parent, but the interjection of the copulating parents which lie at the base of character formation in man. There is then a brief, highly speculative, but provocative, discussion of the fact that man owes his culture, his humanity in fact, to the retardation of his development, and to the long period which intervenes between the appearance of a strong instinctive sexuality in the child, and the time at which that child's body and the society in which he lives will permit him to function on a genital level.

It will be readily apparent from this summary how disjunctive the argument is. Nevertheless, although *The Riddle of the Sphinx* will probably not only not afford much illumination to a novice in the field, but indeed is likely to arouse very definite antagonism and to alienate research workers from its hypotheses, this reviewer believes that in what it stands for and in the problem which it raises it is a very important book, and a valuable contribution in the border-line field between the two sciences.

Dr. Roheim raises two theoretical questions: (a) To what degree can we understand the plot or theme of a culture by reference to the recurrent traumatic situations to which the child is submitted in the family situation? (b) How are we to explain the origin of culture? and by implication of a third: (c) If we frame an hypothesis about the origin of culture, what light does that hypothesis throw upon an understanding of the whole problem of the development of personality in culture today? His answer

[4] S. Zuckerman, *The Social Life of Monkeys and Apes* (London: Kegan Paul, Trench, Trubner and Co., Ltd. 1932).

[5] Ray H. Carpenter, "A Field Study of the Behavior and Social Relation of Howling Monkeys," *Comparative Psychology Monographs*, 1934, X, No. 48.

to the first problem is that the situation to which the child is submitted, such as sleeping under the mother, or having its genitals orally attacked by the father, is determinative of the form which the adult personality, and the expression of the adult personality, the culture, will take. Each adult generation, its character having been formed in this infantile situation, will repeat the same treatment of the child; and the culture, with its peculiar rituals, ceremonials, and emphases, will be perpetuated. It is only in the minds of children that new situations arise, "The planting of yams must have developed from a play activity of some kind; children messing about with the roots that were left over from a meal may have been the real inventors" (p. 262). So while Roheim reiterates that the system is self-perpetuating, he suffers from the handicap which is apparent in so much work which stems from the life-history approach; he sees the plot of human life as beginning, in every sense, with the birth of a new generation, and in his final argument ignores the society into which the first group of children who suffer from this trauma are born.

This discrepancy he tries to solve, not by attacking it in a direct and logical fashion in terms of concrete social situations, but by going on immediately to (b): What is the origin of culture? His argument runs: If we can solve the nature of human culture, and refer it to an infantile situation against which the development of culture is a defense mechanism, then we have proved that the origins of the specific character of any given culture is also to be referred to the infantile situation. In his attempt to answer (b), Roheim stresses the significance of infantile sexuality, and the need for the development of cultural forms which would permit the acting out in ritualized social form of the individual infantile conflict. He then proceeds to make the further assumptions: (1) that the human infant is capable of highly charged sexual feeling which corresponds to an earlier period in the development of human beings when the human individual was ready for mating at a much earlier age; (2) that the need for suppressing this early sexuality until the individual is physically and socially mature enough for the reproductive process lies at the basis of the psychic structure of the individual; (3) that culture originated in the struggle between the sons and the father for control, during which struggle cultural forms were developed which symbolically expressed the early infantile wishes and fears of each present generation and also the guilt arising from the actual aggression of the adult sons in previous generations; (4) culture at any stage in which it is found contains these two elements, a ritual expression of the type of psychic conflict engendered by the peculiar character of the child's early training at the hands of adults, and a survival in ritual form of the adult social conflicts of generations which existed ages ago, in a pre-human past.

In this argument Roheim has based his reasoning upon a strict analogy between the development of the individual and the development of the race. His argument is not so much that ontogeny recapitulates phylogeny as that the development of the human race and its cultural institutions can be derived from a study of the development of the character of a given human individual. He has left out of consideration the nature of society. It is quite plausible to argue that the origin of society lay in a struggle between the old males and the young males within a horde, and that the young males discovered that by combination, i.e., by a first form of social organization, they could overcome the old males, or at least limit their powers and obtain wives for

themselves at an earlier age than had formerly been the case. If such an event took place, however, it would, if we are to apply what we now know about the nature of society, leave its mark upon subsequent forms of social organization. From it would emerge such institutions as clans, age societies, primogeniture, various formulated ways of behavior which made it possible for males to continue to coöperate in a constructive social policy. Each subsequent generation would be utilizing the social devices of preceding generations, not re-enacting the guilt of preceding generations. Now it is quite possible that, as these social forms developed within which group action was possible, there would also develop rituals which expressed in social form the conflicts which were inherent in the family situation, and the lag between infantile potential sexuality and the actualization of that sexuality on an adult level. Thus at any given period the culture forms of a given people must be regarded from at least two points of view: what functions do they serve in holding the society together, and what functions do they serve in permitting the adequate psychic functioning of the constituent individuals who have, during their childhood, been subjected to a type of training which may justifiably be regarded as traumatic?

If one neglects to consider that societies are after all aggregations of individuals which have their own laws of integration, laws which may possibly be paralleled with the laws of integration of an organism, but which are not comparable to the laws which govern the psyche alone, the result is Roheim's theoretical statement. An illustration may make this clearer. Let us take for instance the institution of incest with its world-wide provenience. If we consider it from Roheim's point of view, we regard it as (a) the result of the repression of the infantile sexual wishes toward the parent, and (b) the reflection of an original intra-horde conflict in the remote past. If, however, we take the position that every social phenomenon must be regarded from the standpoint of the development of the individual *and* from the standpoint of the development of society, we would say that the institution of incest as found in any given society today is the result of a synthesis between (a) the repressions which are induced in each young child by the need to keep in abeyance its infantile sexuality until such time as it is physically and socially ready for reproductive activity and (b) the value to the society of the regulation of conflict within the family group, so as to make intra-familial coöperation possible,[6] and that the extensions of these incest taboos beyond the borders of the domestic family are a social device for increasing the size of coöperative groups.[7] Where such institutions as incest, that is institutions with a universal provenience are found, it is reasonable to look for this type of double significance, to construct hypotheses which flow from the biological nature of individual development as well as from the needs of society. It is also possible to find institutions which apparently subserve either the needs of the individuals or the needs of a certain kind of society to a greater or less degree; as, for instance, the segmentation or stratification of society may be a social form which is essential to the integration of individuals under certain types of social conditions, which has no necessary correlate in the nature of individual development; or myths which continue to exist without any

[6] Brenda Seligman, "Incest and Descent, Their Influence on Social Organization," *Journal of the Royal Anthropological Institute*, 1929, LIX, 231-272.

[7] R. F. Fortune, "Incest," *Encyclopedia of the Social Sciences* (The Macmillan Co., 1932).

immediate function in integrating society.

Or let us take another set of social institutions to which the Freudian school have directed their attention, institutions surrounding death which are characterized by a ritualized ambivalence expressing grief on the one hand, and fear which is a derivative of guilt from a repressed hostility towards the individual dead on the other hand.[8]

If we wish to explain the psychic mechanism of ambivalence which informs these rituals and is subserved by them, we can do so by recourse to the Freudian theory of infantile sexuality, and the development of the unconscious. If this be the only order of explanation invoked, the problem of the ready-made existence of the rituals, often highly elaborate in nature, is left unsolved. The Freudian would refer it to a series of historical events which are analogous to the psychic conflicts which are reënacted today, but such a comparative sociologist as Professor A. R. Radcliffe-Brown would instead add the consideration of a separate set of determinants, and would discuss the disphoria induced in a group by the death of one of its members, the disturbance in the social equilibrium and the value of the ritual observances at death in tiding the group over the emergency and re-establishing social equilibrium. The analogy between the conflicts of the individual and the postulated social situations which occurred in the past and are given ritual expression would be abandoned. When the student of society then attempts finally to discuss the whole problem of society and the diverse cultural forms which obtain among different peoples, both types of explanations will be invoked.

The significance of the conflicts of adults, the discrepancies between the type of character structure induced in the childhood of a given group of adults and the particular social situations in which they later find themselves, as a formative factor in culture can, however, be revived in another form. We may go back to Roheim's own very vivid documentation of the differences between contemporary primitive societies and concentrate upon these differences, instead of obscuring them beneath the blanket conception of the *primitive* which is necessary to the argument as it stands.[9] We then find very great differences between the central traumas of different cultures, and the complex and delicate rituals which express these conflicts. While Roheim is willing to lay the origin of new cultural inventions to children, except when he is discussing the conflict between the fathers and sons in the primal horde, it is quite reasonable to attribute the development of these diverse cultural forms to the gradual contribution which adults make to a changing situation in which they have to utilize existing rituals to express infantile conflicts which are of a different order than the conflicts originally expressed in them. The inter-action between the character structure which is produced by a given type of upbringing, the ritual which exists at a given period, and changing social forms—forms which change in response to clues inherent in the social structure itself and to stimuli received from the physical environment, technological inventions and contacts with other peoples—may well be invoked to explain the distinctive character, the *configuration*[10]

[8] See M. Mead, "An Ethnologist's Footnote to Totem and Tabu," *The Psychoanalytic Review*, 1930, XVII, 297-304; and M. Opler, "An Interpretation of Ambivalence in Two American Tribes," to appear in the November number of *Journal of Social Psychology*.

[9] It is methodologically significant that in his final discussion Roheim relies almost entirely upon his Australian material as the ideal primitive, especially in his discussion of how much happier primitive man is than civilized man, a conclusion which is not supported when his Dobuan material is also borne in mind.

[10] Ruth Benedict, *Patterns of Culture* (Boston, Houghton Mifflin and Co., 1934).

of a given culture over time. The hypothesis of a survival of a series of primordial conflicts in present institutions not only does not reinforce, but it actually weakens, Roheim's argument for the close correspondence between the family situation and the content of the ritual of *different* cultures today. The student will find, however, his analyses of these correspondences very valuable and unique in ethnological research.

We may now return to the third problem raised by Roheim's book, the relevance of his hypotheses to an understanding of the relationship between the development of personality and culture today. Particularly illuminating are his exposition of the significance of the lag in individual maturation and his detailed illustration of the way in which the psychic conflicts engendered in this process may be expressed in myth and ritual, and his material upon the way in which the symbolism of myth and ritual may be again used in the idiosyncratic solutions of individuals. The analysis of the plot or theme of any homogeneous culture—if the spurious historical analogies are abandoned we may substitute the word *homogeneous* for the word *primitive*[11]—with direct reference to the type of infantile situation characteristic of that culture, opens up a new approach in the study of culture and of personality. In my last field work, for instance, I was greatly indebted to insights derived from the lectures[12] which Dr. Roheim gave in New York on his return from the field. Furthermore, the hypothesis which Dr. Roheim presents concerning the biological aspects of retardation and their psychic concomitants is provocative and fruitful. The attempt to incorporate in this approach the survival of past social conflicts on an adult level, which makes a period in the evolution of society strictly analogous to a period in the development of a single individual, merely confuses this issue and adds no useful hypothesis. It is necessary to recognize that because society is composed of individuals with a definite biological constitution, and because some social forms, especially the family, are so intimately linked with the biological nature of man that they are virtually inevitable, no explanation of the social order which does not take these factors into account is valid. But it is equally necessary to recognize that society is composed of a *group* of individuals, and that the perpetuation of social forms follows different laws than those of individual consciousness, and social institutions subserve not only the survival of the individual but the survival of the group. Roheim states that civilization is a neurosis and so reveals his fundamental confusion on this point. Individuals may be neurotic, cultural forms may embody a symbolic solution of conflict which they have induced in individuals, but a society, although it may suffer from diseases which are analogous to the diseases which affect a living organism, i.e., defects in structure and disturbances in functioning, is not of the same order as the individual psyche, and if its present forms are to be referred to conflict situations in the remote past, these forms must be regarded as inventions of adaptive mechanisms for group living, not merely as the symbolic expression of the psychic conflict of the adults who participated in any given social conflict.

MARGARET MEAD.
American Museum of Natural History.

[11] See M. Mead, "The Use of Primitive Material in the Study of Personality," *Character and Personality*, 1934, III, 3-16.

[12] Subsequently published in the Roheim Australian Research Number of the *International Journal of Psycho-Analysis*, Jan.-April, 1932.

A DYNAMIC THEORY OF PERSONALITY. By Kurt Lewin. New York: McGraw-Hill, 1935. Pp. ix, 286. $3.00.

This important volume is a collection of originally independent articles which were written at different times and for different occasions. The selection was made for the purpose of giving a comprehensive survey of the theory, methods, experimental results and application of researches done under Professor Lewin's direction. The difficulties which inhere to some extent in collected works, namely, lack of continuity and repetition, are minimal in the present case. The translation is excellent.

Lewin sets for himself the problem of solving for B the equation $B = f(PE)$, where B is the behavior at a particular moment, P the momentary state of the individual, and E the momentary situation in the psychological environment. With this statement of the general problem of psychology there will be little disagreement, though as to the possibility of arriving at a complete solution in any but the simpler perceptual and memory functions there may be considerable doubt. Lewin is fully cognizant of the "weary skepticism" upon this point, but in the face of it he is unequivocal. He believes that it is possible even in the case of "such vital problems as those with which psychoanalysis is concerned."

To accomplish this desirable end it is necessary, in the first place, to abandon the Aristotelian methods which are still characteristic of psychological research in the fields of intelligence, will, affect, etc. These methods are limited to an essentially circular type of explanation. Events are defined in terms of the classes in which they occur, and the classes, in turn, are defined as the sum total of those characteristics which a group has in common. Thus investigation may find that negativism is quite characteristic of three-year-olds, and this negativistic stage is then used as an explanation for the negativism of particular three-year-olds.

The Galileian methodology of physics, on the other hand, ignores classifications and proceeds to the complete description of the unique object and its situation. This complete description constitutes the law in which the constants for the particular situation are substituted for the variables of the generalized law.

For the complete description of the psychological person and the psychological environment new dynamical concepts are required. The person may be conceived to be surrounded by his psychological life-space. This life-space includes regions of free movement and regions which, though existing for him psychologically, are inaccessible by reason of the social situation or because of limitations of his own intellectual or physical powers. Within this psychological life-space are physical and social facts and mental and physical tasks or activities. The latter may have the character of regions which may be entered or vacated. These objects and areas of the psychological environment are all endowed with valences which usually derive from their direct or indirect relations to the satisfaction of needs. Valences differ as to sign and strength. Depending upon these valences are forces which have the properties of strength, direction, and point of application. The first two properties of forces are to be represented through the mathematical concept of the vector. A force is indicated by a "psychobiological" locomotion. "In any case of locomotion there exists a resultant of forces in its direction." The strength of a valence and the direction it imparts to the behavior depends upon the momentary state of the individual's needs, the position of the valence relative to the individual, and

the presence or absence of other valences. One particularly important class of objects includes barriers and boundaries. Barriers prevent direct movement toward or away from goals, necessitating detour behavior. Boundaries may completely surround goals, preventing their attainment, or they may surround persons and thus also prevent attainment of positive valences or escape from negative ones. The perceptual structuring of the field is of the utmost importance for the solution of detour problems. A field of force indicates which force would exist at each point in the field if the individual were at that point. Besides these fairly simple concepts there are more complex ones which may only be mentioned here: conflicts, states of tension, induced valences, and planes of relative unreality.

Approximately the second half of the book is devoted to illustrations of these concepts. This portion includes a consideration of reward and punishment as a means of encouraging or suppressing definite sorts of behavior, substitute behavior, an interpretation of feeblemindedness, and a brief survey of the experimental investigations which Professor Lewin has inspired. It is impossible in the space available to review these papers individually, for the arguments are detailed and the necessary diagrams are numerous. It should be emphasized, however, than this part of the volume is not less important than the earlier chapters, for it constitutes the testing ground for Professor Lewin's theoretical viewpoints. Whatever may be the fate of Professor Lewin's theory, the experimental work here presented will have a permanent place in psychology. It is to be regretted that the final chapter which constitutes a brief survey of these experiments is not longer and more detailed, as this would have added considerably to the value of the volume.

A critical estimate of Professor Lewin's contributions is perhaps hardly warranted at the present time; it will take much study and investigation to evaluate them properly. However, as to their importance for present-day workers there can be no question. They are rapidly assuming a high valence in the "life-space" of psychology. As has already been stated, the experiments which they have inspired are certainly among the most significant of recent years and amply justify the theoretical structure, entirely aside from its eventual fate.

ROGER G. BARKER
LEWIS M. TERMAN.

PRINCIPLES OF GESTALT PSYCHOLOGY. By K. Koffka. London: Kegan Paul, Trench, Trubner and Co. New York: Harcourt, Brace and Co., 1935. Pp. xi, 720. 25/-net. $5.00

For sometime past it has been common knowledge among psychologists that Professor Koffka was engaged on a systematic treatise on Gestalt psychology. Few, however, can have expected so thorough and so far-reaching an exposition. Certainly it will no longer be possible for critics to dismiss Gestalt psychology merely as a one-sided school whose contributions, if they have any value, have been anticipated by others.

The book is not a general survey of the field which can be recommended as a manual to the ordinary elementary student of psychology; still less is it a popular outline suitable for the layman. Though it includes valuable abstracts of a great deal of unfamiliar German experimental literature, it assumes in the reader a rather thorough knowledge of Köhler's *Gestalt Psychology* and his *Mentality of Apes,* and some acquaintance with the previous work of Wertheimer, Koffka, and Lewin. Its main object

is to establish certain fundamental principles on a firm experimental basis, and then to develop these principles consistently so as to show their application not merely to perception and learning but to psychology as a whole including thinking, feeling and action, and the Ego or personality, also to social psychology and even to the philosophy of nature, life, and mind.

We cannot undertake here a full account or criticism of these principles; in brief schematic outline, however, they include the following: (1) The view that psychological and physiological phenomena are ordered or integrated, not the products of mere mechanical interplay or aggregation. Since the same principle applies to many physical phenomena (e. g., crystal structure), it is claimed that the vitalistic-mechanistic dilemma and the mind-body duality are avoided. (2) The concept of organization. The author seldom uses the term Gestalt; instead he clarifies the concept of organization by showing the concrete forms which it takes, e. g., the laws of prägnanz, continuity, proximity, closure, etc. (3) Isomorphism, or the essential similarity of psychological processes and of the underlying physiological processes. Though this principle finds its chief application in the discussion of the perception of motion, it is of course a corollary of (1), and it is basic to the next item. (4) The reality of shape, both as a direct datum of experience, not built up out of a mosaic of elementary sensations and associations, and as a physiological unit which cannot be reduced to the traditional neuronic arcs; in other words, the "meaning theory" is rejected even more emphatically than in Köhler's book. This principle is applied mainly to the visual perception of simple figures and of such complex shapes as "a gloomy landscape" or "an irritable individual." It is unfortunate, perhaps, that more attention is not paid to auditory and other sense fields, for a modified form of meaning theory is less easily disposed of there than in vision. Surely no psychologist nowadays believes (as the author seems to assume) that mere repetition of contiguous sensations is sufficient to establish the percept of a shape; whereas a great many psychologists consider that the segregation of an object is dependent on a conative reaction to the object, so that therefore acquired experience *is* essential to a good deal of perceptual organization. Indeed we suspect that this modification of the meaning theory is reintroduced into Gestalt psychology by the author himself under the guise of the next principle. (5) The importance of the Ego as the source of "attention" and attitudes, emotions, and needs. At the same time it is itself an organized and segregated object, similar to its own perceptions, memories, etc. Perhaps the most outstanding feature of the book is this *rapprochement* that it offers between the neo-structuralist and the functionalist viewpoints. But here the discussion is disappointingly restricted, for the sound reason that there are at present so few experimental facts to build on. Research on the psychology of personality and on social psychology in the light of Professor Koffka's suggestions is most urgently needed.

There is of course far more material than we have mentioned which claims the attention both of supporters and of critics (for example, over two hundred pages on memory and learning). In general the author's handling of so large and complex a field is remarkably clear and fluent, and he is less provocative and dogmatic than most previous writers pro or con. Particularly noteworthy are his capacity for marshalling experimental evidence and his recognition

of lacunae both in experiment and theory. With few exceptions (the worst of which is the transposition of several lines near the bottom of p. 371), the printing is excellent, and the numerous figures in the text are adequate.

P. E. VERNON.

THE SUBNORMAL MIND. By Cyril Burt. London: Humphrey Milford, The Oxford University Press, 1935. Pp. viii, 368. 10*s.* 6*d.*

It is ten years since Professor Burt promised, in his preface to *The Young Delinquent*, two further volumes on the subnormal school child. What is presumably one of these volumes, entitled *The Backward Child*, is now in the press. But the book under review seems to be a summary of the whole trilogy, for it deals not only with the delinquent and the defective but also with the neurotic. Since it is based on the author's Heath Clark lectures at the University of London in 1933, it is addressed primarily to medical officers of health; but the reviewer has no hesitation in recommending it to all who work in child guidance clinics, whether doctors, psychologists or social workers, to teachers and others who are practically concerned with child psychology. The earlier chapters are a simple and very readable outline of the ground which the author covered so fully in his *Mental and Scholastic Tests* and *The Young Delinquent*. But they include many fresh points. For instance, the influence of intelligence, as tested in childhood, on adult level of employment is strikingly demonstrated; the careers of a large number of dull and backward children have been followed up (not certified defectives, but school children with I.Q.'s below 85), and after 10 to 15 years it was found that 21% of them were unemployed, 25% engaged on unskilled labor, 38% on semi-skilled, and only 16% on work "which might be classified as skilled, but was of an exceedingly limited and wholly routine or mechanical order." Even more significant, perhaps, are some figures on recidivism: Professor Burt finds that 70% of delinquent children who were referred to him before the age of 9 may now "fairly be described as cured"; of cases referred between 9-14 years only 57%, of adolescents 62%, and of adults only 32% have not returned to crime. It is impossible in this review to mention innumerable other matters of practical and theoretical interest with which the book deals.

The later chapters on the neurotic child are both the most original and the most detailed. The reviewer cannot call to mind a better account of the neuroses of childhood. The author points out that, while approximately 1% of British children are mentally defective, probably 5% are suffering from neurotic disorder so serious as to need immediate special treatment. A far greater proportion must be regarded as "emotionally subnormal," for there is no sharp dividing line between neurosis and the minor anxieties and behavior problems of the so-called normal child. The symptoms, causation, incidence, treatment and theory of different types of neurosis are fully discussed.

Finally, there are three useful appendices. The first reproduces some of Professor Burt's graded tests of educational attainments (reading, spelling, and oral arithmetic). The second lists the age assignments of the tests in the Stanford-Binet scale according to the author's English restandardization. This revision is used by most child guidance clinics in the country, but has never been published. The third is a questionnaire on neurotic symptoms, along the lines of the Woodworth inven-

tory and Thurstone's personality schedule; its limitations are carefully noted.

If there is one section of the book which seems relatively weak or open to criticism, it is Professor Burt's theory of temperament. The cornerstones of the theory are still the list of about a dozen discrete instincts, and the conceptions of "general emotionality" and of "stable-unstable" and "repressed-unrepressed" types. When first put forward twenty years ago this schema was a brilliant extension of McDougall's *Introduction to Social Psychology*, but much has happened in the interval. Professor McDougall's views on instincts have undergone modification, experiments with personality tests have cast grave doubts on the reality of a general emotionality factor, and, above all, the work of the psychoanalysts has demonstrated a far greater complexity in the organization of personality than can be accounted for by instincts and sentiment formations. Professor Burt is fully cognizant of these developments. Hence his insistence on the determination of a subnormal child's innate emotionality or instability, as well as his or her intelligence, and his attempt to trace neurasthenia down to "the instinct of repulsion," etc. appear somewhat out of touch with present-day theory and practice. In reply to this criticism it might be pointed out that no other theory of the foundations of temperament and character, nor systematic classification of emotional types, has won general acceptance. Professor Burt has certainly provided us with a sound and well-balanced account of the contributions of social and applied psychology, of statistical investigation and of medical psychology to the study of the subnormal child, even if he has not quite succeeded in welding them into a single integrated viewpoint.

P. E. VERNON.

Books Received

IMAGINATION IN EARLY CHILDHOOD AND ITS FUNCTION IN MENTAL DEVELOPMENT. By Ruth Griffiths. Preface by J. C. Flugel. London, Kegan Paul, Trench, Trubner, 1935. 367 pp. with index. 12/6.

SCIENCE AND THE HUMAN TEMPERAMENT. By Erwin Schroedinger. Translated and with a Biographi-Introduction by James Murphy. Foreword by Lord Rutheford of Nelson. London, G. Allen & Unwin, 1935. 154 pp. 7/6.

INDIVIDUAL HEALTH: *A Technique for the Study of Individual Constitution and Its Application to Health.* By E. Obermer. Vol. I. Biochemical Technique. By E. Obermer and R. Milton. London, Chapman and Hall, 1935. 244 pp. with index. 15/-.

GEISTESKRANKHEIT UND HANDSCHRIFT. *(Zeitschr. f. d. gesamte Neurologie und Psychiatrie.)* By Heinrich Unger. Berlin, Julius Springer, 1935. Pp. 569 to 604.

AWARENESS AND THE NEUROSES OF DECLINING YEARS. By M. Ferguson, H. Weber, O. H. Woodcock, J. C. Young. Individual Psychology Medical Pamphlets No. 14. London, C. W. Daniel Co., 1935. 61 pp. 2/6.

WRITING AND WRITING PATTERNS: *In Two Sets of Hinged Cards A and B: Five Books of Copies and a Teacher's Book.* By Marion Richardson. London: University of London Press. 5/-.

ABILITY AND KNOWLEDGE: *The Standpoint of the London School.* By Frank C. Thomas. With a Foreword by C. Spearman and an Appendix by E. Walters. London, Macmillan, 1935. 388 pp. with index. 15/-.

THE CHEMISTRY OF THOUGHT. By C. A. Claremont. London, G. Allen & Unwin, 1935. 259 pp. with index. 8/6.

HUMAN PERSONALITY AND ITS SURVIVAL OF BODILY DEATH. By F. W. H. Myers. The Swan Library, Vol. 28. London, New York, Toronto: Longmans, Green, 1935. 307 pp. with index. 3/6.

PSYCHOLOGY: A FACTUAL TEXTBOOK. By E. G. Boring, H. S. Langfeld, H. P. Weld, and Collaborators. New York, John Wiley and Sons; London, Chapman and Hall, Ltd., 1935. xviii, 555 pp. $2.75.

FACTOR ANALYSIS IN THE STUDY OF PERSONALITY. By John Clemans Flanagan. Stanford University, Calif., Stanford University Press, 1935. x, 103 pp. $1.25.

MENTAL HEALTH: *Past, Present, and Future.* By Arthur Hiler Ruggles. Baltimore, Md., Williams and Wilkins Co., 1934. 104 pp.

PACEMAKERS IN RELATION TO ASPECTS OF BEHAVIOR. By Hudson Hoagland. New York, Macmillan Co., 1935. x, 138 pp. $3.00.

THE FOUNDATIONS OF HUMAN NATURE: *The Study of the Person.* By John Morris Dorsey. New York, Longmans, Green and Co., 1935. xiii, 488 pp. $2.80.

THE PROBLEM OF MENTAL DISORDER. Edited by Madison Bentley and E. V. Cowdry. New York and London, McGraw Hill Book Co., 1934. x, 388 pp. $4.00.

EMOTIONS AND BODILY CHANGES. By H. Flanders Dunbar. New York, Columbia University Press, 1935. ix, 595 pp. $5.00.

PERSONALITY AND THE FAMILY. By Hornell Hart and Ella B. Hart. New York and London, D. C. Heath and Co., 1935. xiii, 381 pp. $2.80.

I. L. PERETZ, PSYCHOLOGIST OF LITERATURE. By A. A. Roback. Cambridge, Mass., Sci-Art Publishers, 1935. 300 pp. $2.65.

WHY WE FEEL THAT WAY: *An Analysis of the Human Emotions.* By August W. Trettien. Boston, The Stratford Co., 1935. vi, 452 pp.

THE PSYCHOLOGY OF WANTS, INTERESTS AND ATTITUDES. By Edward L. Thorndike. New York and London, D. Appleton-Century Co., 1934. x, 301 pp. $3.50.